Time-delay Cosmography with New Angular Diameter Distance Measurements

Inh Jee

München 2017

Time-delay Cosmography with New Angular Diameter Distance Measurements

Inh Jee

Dissertation
an der Astronomie
der Ludwig–Maximilians–Universität
München

vorgelegt von
Inh Jee
aus Seoul, Süd Korea

München, den 13 März 2017

Erstgutachter: Prof. Dr. Eiichiro Komatsu
Zweitgutachter: Prof. Dr. Sherry Suyu
Tag der mündlichen Prüfung: 9. Mai 2017

Contents

List of Figures

List of Tables

Zusammenfassung

Wir präsentieren eine neue Methode zur Entfernungsmessung im Universum unter Verwendung der zeitverzögerten Ankunftszeiten des Lichts einer variablen Quelle, das durch eine Galaxie im Vordergrund stark gebeugt wird. Die allgemeine Relativitätstheorie postuliert, dass Masse die Raum-Zeit verzerrt. Somit wird die Trajektorie von Licht gebeugt wenn es ein Objekt passiert das als Gravitationslinse agiert. Diese Verzerrung verursacht Abweichungen in den Pfadlängen der Photonen, welche sich in unterschiedliche Richtungen ausbreiten. Die Photonen wandern auf ihrem Pfad um die Gravitationslinse herum auch durch unterschiedliche Gravitationspotentiale und erfahren eine Zeitdilatation. Dadurch unterscheiden sich die Ankunftszeiten der Photonen beim Beobachter, auch allgemein bekannt als Zeitverzögerung. Die Zeitverzögerung ermöglicht somit einen Rückschluss auf das entlang der Sichtlinie integrierte Gravitationspotential der Gravitationslinse, in der Größenordnung von $GM\ln(r)$. Demgegenber steht die ungeordnete Bewegung der Sterne welche der Gravitationsanziehung der (Linsen-)Galaxie entgegenwirkt und mit welcher sich das Gravitationspotential, GM/r, der Gravitationslinse ebenfalls, durch die Messung der stellaren Geschwindigkeitsdispersion, abschätzen lässt. Mit diesen beiden Informationen der Gravitationslinse lässt sich die Größe r des Systems ableiten und durch den Vergleich der physikalischen Größe r und der Winkelausdehnung θ, gemessen anhand der Abbildungspositionen, die Winkelausdehnungsentfernung zur Linse ermitteln.

Die ursprüngliche Idee der Verknüpfung von Zeitverzögerungen und stellaren Geschwindigkeitsdispersionen zur Messung der Winkelausdehnungsentfernung wurde 2009 von Paraficz und Hjorth präsentiert. Nach einer Einführung in die Grundlagen des Gravitationslinseneffekts erweitere ich ihre ursprüngliche Abschätzung, basierend auf einfachsten Linsenmodellen, um Massenprofile welche beliebigen Potenzgesetzen folgen können und eine anisotrope Geschwindigkeitsstruktur besitzen. Wir bestätigen, dass die Distanzen auch mit realistischeren Gravitationslinsensystemen abgeschätzt werden können. Eine bemerkenswerte neue Erkenntnis hierbei ist, dass bei der Berechnung der Winkelausdehnungsentfernung die Entartung der Linsenmasse, aufgrund der Existenz von Masse entlang der Sichtlinie, sich aufhebt, welche eine der gravierendsten Einschränkungen für die Nutzung des Gravitationslinseneffektes als kosmologische Probe darstellt. Wir wenden diese Methode an zwei Linsen an und messen somit zum ersten mal ihre Winkelausdehnungsentfernung mit einer Genauigkeit von 13-18%.

Die gemessene Winkelausdehnungsentfernung ist invers proportional zur radialen Komponente der stellaren Geschwindigkeitsdispersion, die aufgrund der Anisotropie der Ge-

schwindigkeitsstruktur jedoch nicht direkt beobachtbar ist. Demnach spielt die Kinematik der Linsengalaxie eine wichtige Rolle für die erfolgreichen Anwendung unserer Methode. Wir untersuchen den Einfluss der Geschwindigkeitsanisotropie anhand zweier Modelle, basierend auf Osipkov und Merrit (1979 und 1985) und ihrer zwei-Parameter Erweiterung. Wir wenden die sphärische Jeans Gleichung mit diesen parametrisierten Anisotropiemodellen an um die radiale Geschwindigkeitsdispersion zu berechnen, und projizieren sie entlang der Sichtlinie um eine beobachtbare Dispersion zu ermitteln. Wir finden heraus, dass das Verhältnis der Aperturgröße zum effektiven Radius der Linse die Messfehler in der vorhergesagten Geschwindigkeitsdispersion beeinflusst. Wir zeigen auf, dass $R_{ap}/R_{eff} > 0.7$ notwendig ist um die von der Anisotropie erzeugten Fehler geringer zu halten als die momentanen Messunsicherheiten. Vorausgesetzt dass die Geschwindigkeitsdispersion an einem Radius gemessen werden kann an welchem der Einfluss der Anisotropie minimiert wird (Idealpunkt), lässt sich die Unsicherheit in der Winkelausdehnungsentfernung aufgrund der Anisotropie bis auf 12-13% verringern.

Ich zeige, dass Einschränkungen bezüglich der Kosmologie die mithilfe der Winkelausdehnungsentfernung von Gravitationslinsen getroffen werden mit anderen Distanzproben konkurrenzfähig sind. Unter Verwendung realistischer Abschätzungen für die Quell- und Linsenverteilungen im bevorstehenden Large Synoptic Survey Teleskop (LSST) berechnen wir die erwarteten Einschränkungen für die Kosmologie und vergleichen sie mit jenen anderer Proben wie z.B. Supernovae (SNe) Typ Ia, der kosmologischen Hintergrundstrahlung (CMB) und baryonisch akustischen Oszillationen (BAO). Wir finden heraus, dass mit einer 5% (10%) Messgenauigkeit beider Linsensdistanzen, das heißt der Zeitdilatationsentfernung und der Winkelausdehnungsentfernung je System, die momentanen BAO+CMB+SNe Vorhersagen bezüglich der zeitveränderlichen dunklen Energie Zustandsgleichung um einen Faktor zwei (20%) verbessert werden. Im Vergleich zum Fall ohne Gravitationslinse lassen sich die Vorhersagen bezüglich der Gütezahl der dunklen Energie hierbei ebenfalls um einen Faktor zwei (50%) verbessern. Wir zeigen, dass diese signifikante Verbesserung dem Gravitationslinseneffekt zuzuschreiben ist, welcher die Entartung zwischen der Krümmung und den Parametern der Zustandsgleichung aufzuheben vermag.

Zusammenfassend erläutere ich die Aussichten und Herausforderungen für die Beobachtung zeitverzögerter Gravitationslinsen. Untersuchungen, welche ein großes Volumen am Himmel in der Zeitdomäne abdecken, werden in der nahen Zukunft zur Verfügung stehen und es wird erwartet dass diese Untersuchungen die Anzahl der beobachteten zeitverzögerten Gravitationslinsen auf mehrere hundert anwachsen lässt. Die aufgrund starker Gravitationslinsen erzeugten Bilder im beobachteten Volumen zu finden stellt jedoch weiterhin eine große Herausforderung dar. Des Weiteren erfordert die Messung der Zeitverzögerung zwischen den Bildern und der Geschwindigkeitsdispersion der Gravitationslinse eine speziell dafür vorgesehene Nachuntersuchung; hohe Kadenzüberwachung sowie hoch aufgelöste Spektroskopie sind notwendig um Massenmodelle der Gravitationslinse erstellen zu können. Ich erörtere den momentanen Stand und mögliche Lösungen zu diesen Herausforderungen.

Abstract

We present a new method to measure distances in the Universe using delays of arrival times of lights of a variable source that is strongly lensed by a foreground galaxy. General theory of relativity states that mass distorts space-time and thus bends the path of light as it passes by an object which acts as a lens. This distortion creates differences between the path lengths of photons propagating in different directions. These photons also go through different gravitational potentials along the paths around the lens, and experience time dilation. Thus the arrival times of the photons differ, which is called time delay. The time delay thus gives a gravitational potential of the lens integrated along the line of sight, giving approximately $GM\ln(r)$. On the other hand, as the random motion of individual stars counteracts the gravitational attraction of the lens galaxy, one can estimate the gravitational potential of the lens, GM/r, by measuring the stellar velocity dispersion. With these two pieces of information of the lens, one can deduce the size of the system, r, and by comparing the physical size of the lens r to the angular size θ measured from the image positions, the angular diameter distance to the lens can be constrained.

The original idea of combining time delays and stellar velocity dispersions to measure angular diameter distances was presented in 2009 by Paraficz and Hjorth. After introducing the basics of gravitational lensing, in this thesis I first expand their original estimate based on the simplest lens model, to include mass profiles with arbitrary power laws and the velocity structure that is anisotropic. We confirm that the distances can still be estimated from more realistic lens systems. A remarkable new finding is that in calculating the angular diameter distance, the lens mass degeneracy due to the existence of external mass along the line-of-sight cancels, which is one of the major limiting factors for gravitational lenses as a cosmological probe. We apply this method to two lenses and measure their angular diameter distances for the first time, with precision of 13-18%.

The measured angular diameter distance is inversely proportional to the radial component of the stellar velocity dispersion, which is not observable directly due to anisotropy in the velocity structure. Thus the kinematics of lens galaxies plays a crucial role in successful application of our method. We investigate the impact of velocity anisotropy using two models based upon Osipkov and Merritt (1979 and 1985, respectively), and its two-parameter extension. We use the spherical Jeans equation with these parametric anisotropy models to calculate the radial velocity dispersion, then project it along the line-of-sight to compute an observable dispersion. We find that the ratio of the size of the aperture to the effective radius of the lens changes the uncertainty in the predicted velocity dispersion. We show

that $R_{ap}/r_{eff} > 0.7$ is required to keep the uncertainty caused by the anisotropy smaller than the current measurement uncertainty. We also find that assuming that the velocity dispersion can be measured at a radius where the impact of anisotropy is minimized (sweet spot), the uncertainty in angular diameter distance due to anisotropy can be reduced to 12-13%.

I show that the constraints on cosmology from angular diameter distances of lenses is competitive to other distance probes. Using realistic predictions of the source and lens distributions expected from the upcoming Large Synoptic Survey Telescope (LSST), we calculate the expected constraints on cosmology and compare them with those from the other probes such as Supernovae (SNe) type Ia, Cosmic Microwave Background (CMB), and Baryon Acoustic Oscillation (BAO). We find that with 5% (10%) precision measurement of both lensing distances, the time-delay distance and the angular diameter distance per system, would improve the current BAO + CMB + SNe constraints on time-varying dark energy equation of state by a factor of two (20%), and those on the overall figure of merit of dark energy by about a factor of two (50%) relative to the case with no lensing. We show that this significant improvement is due to lensing's ability to break the degeneracy between curvature and the equation of state parameters.

To conclude, I give the future prospects and challenges in observing time-delay lenses. Surveys covering large volumes of the sky in time domain are becoming available in the near future, and it is expected that these surveys will increase the number of observed time-delay lenses to hundreds. However, the problem of finding strongly lensed images in the observed volume still remains a challenge. Also, measuring the time-delay between images and the velocity dispersion of the lens requires dedicated follow-ups, both in monitoring the images with high cadence and in obtaining high resolution spectroscopy to construct the lens mass model. I discuss the current status and possible solutions to these challenges.

Chapter 1

Strong Gravitational Lensing

I start with a brief introduction to strong gravitational lensing. I begin by reviewing the theory, then more to observable quantities, discuss model degeneracies and finally applications to cosmology. The notation defined in this chapter will be used throughout the paper.

1.1 Theory

The general theory of relativity predicts that the existence of mass distorts space time. A light propagates through the curved space-time, its path appears to be bent to the observer, which is known as gravitational lensing. In classical physics, the deflection of light path as they pass through media with different refractive indices has been known and used for a long time, which is where the phenomenon obtained its name "lensing". One of the major differences between gravitational lensing and optical lensing is that gravity is achromatic, so there is no frequency dependence in gravitational lensing, except for specific cases where the light propagates in a plasma [15]. The object which deflects the light is defined as a lens, and the object which emits the light is the source (see schematic diagram in 1.1). As the Universe is full of objects with mass, lensing is a very general phenomenon. However, visual distortion of images is rare as the lens should have sufficiently high surface mass density to cause strong distortion of space-time. Also, the lens and the source should be in close proximity to each other in angular scale.

To state this quantitatively, gravitational lensing can be separated into two regimes, weak and strong, depending on how massive the lens is - more precisely, the surface density of the lens exceeding the critical density satisfies a sufficient condition for the system to be in strong gravitational lensing regime. The relative position of the source and the lens also plays a role: the source should be close enough to the lens to be strongly lensed. The critical density, $\Sigma_{\rm cr}$, is defined as follows:

$$\Sigma_{\rm cr} \equiv \frac{c^2 D_{\rm s}}{4\pi G D_{\rm d} D_{\rm ds}}, \tag{1.1}$$

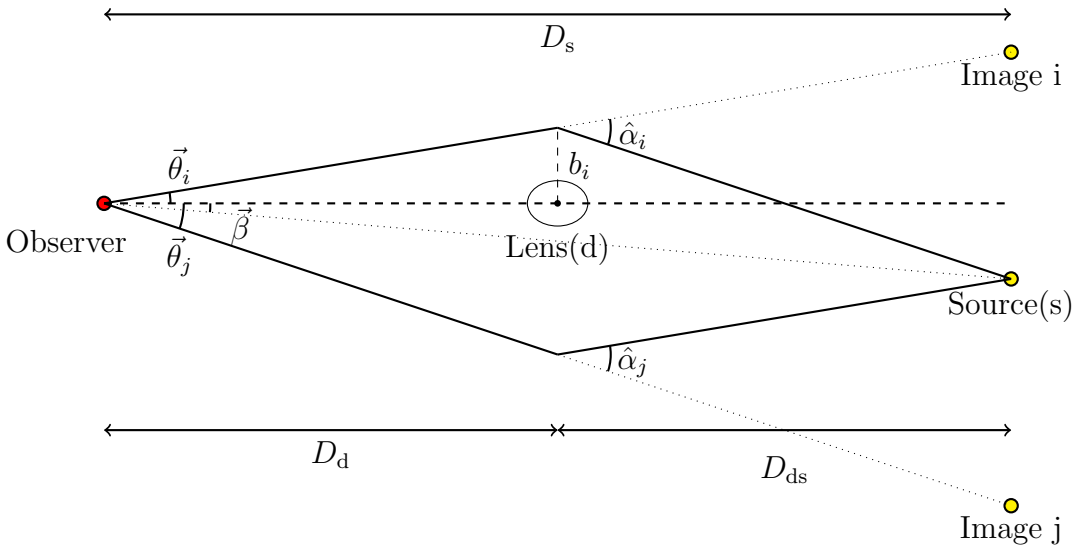

Figure 1.1: Configuration of a strong lens system, with definition of the variables used throughout this thesis. All angles are measured with respect to the center of the lens galaxy; $\vec{\theta}$ is the angular position of the image; $\vec{\beta}$ is the angular position of the source in the absence of the lens; $\vec{\alpha}$ is the scaled deflection angle; $\hat{\alpha}$ is the deflection angle at the lens plane; and $\vec{b}$ is the physical separation to the closest approach at the lens plane.

where D is the angular diameter distance and subscripts d and s stand for the deflector and the source, respectively. When the surface density of the lens exceeds the critical density, and when the alignment between the lens and the source is perfect, the images will be stretched tangentially to form a ring around the lens, which is called an *Einstein ring*. When the surface density of the lens is below the critical value, i.e. in weak lensing regime, an image of a source will be distorted, resulting changes in ellipticity which cannot easily be distinguished from its intrinsic shape. A statistical analysis based on assumptions about the intrinsic shape of galaxies is applied to make use of weak lensing information.

In this thesis we focus on galaxy scale lensing, where both the source and the lens are galaxies. In this case, both the source and the lens appear as extended images in the observation, but for simplicity we consider the case of point sources in this section.

Let the angular position of the image be $\vec{\theta}$ and that of the source be $\vec{\beta}$, as shown in fig. 1.1. The arrival time between photons from different images varies, which produces a *time delay*. The absolute time delay per image can be written as

$$t(\vec{\theta}, \vec{\beta}) = \frac{1}{c}(1 + z_{\rm L})\frac{D_{\rm d} D_{\rm s}}{D_{\rm ds}}\phi(\vec{\theta}, \vec{\beta}), \tag{1.2}$$

where ϕ is the so-called *Fermat potential*, which is defined as

$$\phi(\vec{\theta}, \vec{\beta}) \equiv \frac{(\vec{\theta} - \vec{\beta})^2}{2} - \psi(\vec{\theta}). \tag{1.3}$$

The first and the second terms in equation (1.2) are geometrical and potential time-delay terms, respectively. Here, ψ is the lens potential, which is calculated as

$$\psi(\vec{\theta}) = \frac{1}{\pi} \int d^2\theta' \kappa(\vec{\theta'}) \ln |\vec{\theta} - \vec{\theta'}|, \tag{1.4}$$

where the lensing convergence field, κ, is defined by

$$\kappa(\vec{\theta}) \equiv \frac{\Sigma(\vec{\theta})}{\Sigma_{\rm cr}}. \tag{1.5}$$

The projected surface mass density, Σ, is

$$\Sigma(\vec{\theta}) = \int_{-\infty}^{\infty} \rho[D_{\rm d}\vec{\theta}, \ell] \; d\ell, \tag{1.6}$$

where ℓ denotes the line-of-sight coordinate, and the critical surface mass density is defined in 1.1. Physically, when $\kappa > 1$, the system satisfies the sufficiency condition to form multiple images.

The absolute time delay, t, is not an observable as we cannot directly observe the source without the lens, or the time difference between lensed and un-lensed images. However, if we have multiple images, we can compare the relative time delay between image pairs to calculate the time delay between two (or more) lensed images. Also, ϕ can be modeled to satisfy observational constraints such as image positions, flux ratios and time-delay differences between multiple pairs of images; thus, we can obtain the so-called *time-delay distance*, which I will introduce in section 1.2.1.

In a differential form, the lens potential is related to the convergence field via

$$\kappa(\vec{\theta}) = \frac{1}{2}\nabla^2\psi(\vec{\theta}), \tag{1.7}$$

where ∇ is a derivative in $\vec{\theta}$ coordinates. Now we can write the *lens equation* which relates the observed image position to the source position in terms of the lens potential,

$$\vec{\theta} - \vec{\beta} = \nabla\psi(\vec{\theta}) = \vec{\alpha}, \tag{1.8}$$

where $\vec{\alpha}$ is the scaled deflection angle.

The magnification μ is defined in terms of the image, $\vec{\theta}$, and the source, $\vec{\beta}$, in the following way:

$$\mu = \frac{1}{\det\,(\frac{\partial\vec{\beta}}{\partial\vec{\theta}})} \equiv \frac{1}{\det\,A}, \tag{1.9}$$

and describes the ratio of area between the image and the source. Here, A is the Jacobi matrix that governs the mapping between the image plane and the source plane, or the Hessian matrix of the projected potential ψ, which can be written as

$$A = \begin{pmatrix} 1 - \psi_{,11} & -\psi_{,12} \\ -\psi_{,21} & 1 - \psi_{,22} \end{pmatrix}, \tag{1.10}$$

where $x_{,i} \equiv \frac{\partial x}{\partial \theta_i}$ and thus $\psi_{,12} = \psi_{,21}$. By separating the isotropic and anisotropic components, A can be rewritten as

$$A = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{1.11}$$

where $\kappa \equiv \frac{\psi_{,11}+\psi_{,22}}{2}$ (consistent with equation 1.7), $\gamma_1 \equiv \frac{\psi_{,11}-\psi_{,22}}{2}$ and $\gamma_2 \equiv \psi_{,12}$. By construction, the convergence κ only changes the size of images isotropically, while the shear components γ_1 and γ_2 distort the shape of images. The eigenvalues of A are

$$\lambda_\pm = 1 - \kappa \pm \sqrt{\gamma_1^2 + \gamma_2^2}, \tag{1.12}$$

thus when $\lambda_\pm = 0$, the magnification μ becomes infinity. The curve where the magnification becomes infinite is the *critical curve* on the image plane, and the critical curve mapped onto the source plane is called the *caustic*. Depending on λ, the line that corresponds to λ_+ is called the radial critical line and λ_- is the tangential critical line. Near the radial critical line the images are distorted perpendicular to the curve, while near the tangential critical line the images are stretched tangentially along the curve. The number of images changes depending on whether the source is located inside (quadruple) or outside (double) the inner caustic, and the image configuration changes depending on where the source is located with respect to the caustic (see Figure 1.2). In galaxy-galaxy strong lens systems, the image configurations are either double ($\sim$75%) or quadruple ($\sim$ 25%) in most cases.

1.1.1 Observables

Strong lensing provides observables which can be used to estimate the mass distribution of the deflector, including its substructures. Most of the observables are based on images: the image positions and their fluxes are observable. Source properties such as position and flux can be inferred, as the images are mapped back to the source plane as the deflector mass is reconstructed, but they are usually not direct observables, as the light of the source are all deflected and forms images instead of reaching the observer unlensed. This adds one degree of freedom, the source position, to the lens mass modeling and causes degeneracies in the model, which will be discussed in section 1.1.2.

When the source is variable, time-delay Δt between images can be measured. Common examples of variable sources are quasars and supernovae. However, the first detection of a lensed supernova was not made until 2014 [85], due to the rarity of these events [72, 78]. There are a handful of lensed quasars which are thoroughly studied, including monitoring over extended periods to measure the time delay between images. One of the major uncertainties in measuring time delay is micro-lensing by individual stars in the lens galaxy. Having very high surface mass density, stars distort critical lines effectively, and as a result the magnification of individual images varies stochastically as they pass through the critical lines. The timescale of this variability is on the order of weeks. Quasars have light curves which are also intrinsically stochastic, thus no light curve templates can be

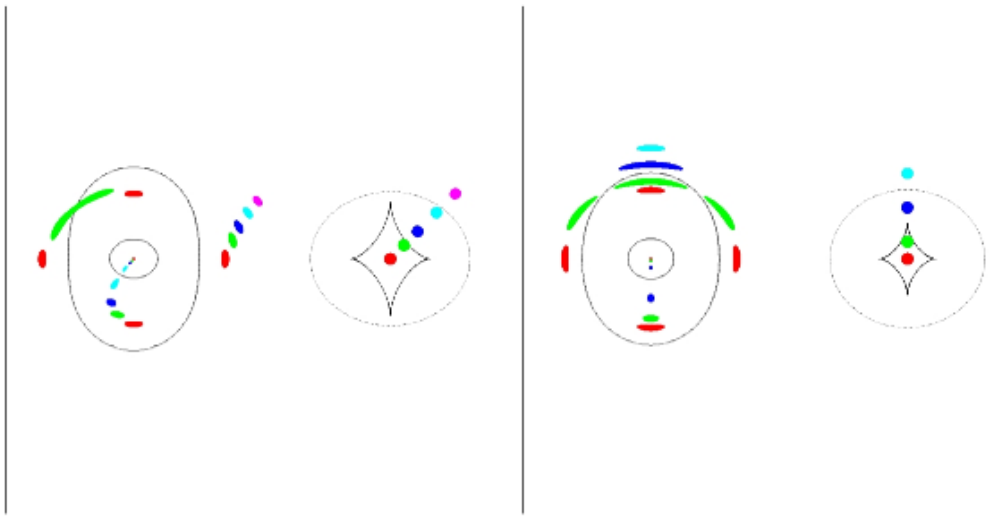

Figure 1.2: Critical curves / image positions (left of each panel) and caustic / source positions (right of each panel) for an elliptical lens and a compact source. Two specific configurations which form as the source reaches the caustic are shown in green: Image configurations shown in the left panel are called *folds*, while the configurations on the right are called *cusps*. When the source is positioned at the center of the caustic (red points), the configuration is called *cross*. In each panel, each color corresponds to the images of its source. The outer (inner) curves in the image planes are the tangential (radial) critical lines, and the outer (inner) curves in the source plane are the radial (tangential) caustics. Figure adopted from [76] and color coded by Massimo Meneghetti.

made. High cadence observation, with period on the order of a day, is required to effectively measure time delay.

In most cases, the original flux of the source is not known (exceptions include type Ia supernovae, where the absolute magnitude of the source is roughly known and can be calibrated from the shape of the light curve and the host galaxy properties), thus the absolute magnification of an image with respect to the source is not known. However, as the brightness of individual images is observable, the magnification ratio between images relative to each other can be determined observationally. Assuming a smooth mass distribution, the magnification ratio is well predicted by the configuration of the image positions, thus deviations from the predicted magnification ratio indicate of anomalies, such as existence of clumpy substructures in the halo [66, 70].

1.1.2 Degeneracies

Multiple degeneracies are present in estimating the lens mass distribution from lensing observables. The fundamental problem arises as three-dimensional properties of the lens are estimated from observables that are projected onto two dimensions. Also, in most

cases the source is out-shined by the lens, so that the shape, position and brightness of the source is not observable. The observables are thus relative properties between images, rather than relations between the source and each image. Thus, direct mapping from the source to the images is not possible. Instead, the multiplicity of the images is used: by imposing a condition that each image should map back to the same source, the validity of the lens mass model can be tested. Degeneracies preserve the image properties, especially ones relative to each other, while modifications of the source properties, mass model and its normalization are allowed. These degeneracies only relate to the effect on lensing observables in global scales, and do not include local substructures that can introduce various additional degeneracies and anomalies. Extra information, other than lensing observables, is required to break these degeneracies.

Lensing conserves the surface brightness between the lens and the source, following photon energy conservation. Thus, magnification is determined by the ratio between the area of the images and that of the source, which is consistent with the definition in equation 1.9. Starting from the lens equation, [31, 39] have categorized the mass model degeneracies into three categories: *Similarity transformation* (ST), *magnification transformation* (MT; later referred to as *Mass-Sheet Transformation* (MST)) and *prismatic transformation* (PT). Later, *Source position transformation* (SPT), which is a more general from of lens convergence transformation that has been introduced in [92].

ST is when all the distances are scaled by the same factor, while the deflection angle and the angular position of the source are not changed. Later, [89] approached these transformations using the time-delay equation, and used a different definition of ST: the transformation scales both geometric and potential time delays by the same factor. According to this definition, the previously defined ST is called *distance degeneracy* (DD). Following this definition, [89] adds another transformation to ST, called *angular degeneracy* (AD). The transformation scales geometric and potential time delays by a factor s, while both image positions and the source position are scaled by $\sqrt{s}$. However, due to the change in image positions, this degeneracy can be broken when the images are resolved.

MST is possible due to the unknown source properties. This transformation leaves the observables (image positions and magnification ratios between images) unchanged, while changes the time-delay, absolute magnification and the source position by a constant factor. Also, while the total time delay is scaled by a constant factor, geometric and potential time delays can be scaled by a different factor. Any family of transformations that follows

$$\kappa'(\vec{\theta}) = s\kappa(\vec{\theta}) + (1 - s) \tag{1.13}$$

are considered as MST, where s is an arbitrary constant. In a physical picture, the transformation involves the addition of a constant convergence $1 - s$, and rescales the lens convergence by s to match the image positions. This changes the deflection angle by the lens, and source position should be scaled by a factor of s to keep the image positions. Because the source position is changed while image positions are invariant, magnification is scaled by s^{-2}. To account for the change in magnification while conserving the surface brightness and the distances, the source luminosity should be scaled by s^{-2}.

Source position transformation (SPT) [92] allows the change in not only the normalization of the lens mass distribution, but also the shape of the mass profile. Lensing is sensitive to enclosed mass inside the Einstein radius, which is only sensitive to the local mass profile at that radius. Thus when lens mass is reconstructed from lensing observables, radii far from the Einstein radius cannot be constrained well, thus SPT is allowed. In terms of lensing equation, SPT can be written as

$$\vec{\alpha}(\vec{\theta}) = \vec{\theta} - \vec{\beta} \rightarrow \vec{\alpha}'(\vec{\theta}) = \vec{\theta} - \vec{\beta}'(\vec{\theta} - \vec{\alpha}), \tag{1.14}$$

and any one-to-one transformation $\vec{\beta} \rightarrow \vec{\beta}'(\vec{\beta})$ that preserves $\vec{\theta_i} - \vec{\alpha}'(\vec{\theta_i}) = \vec{\theta_j} - \vec{\alpha}'(\vec{\theta_j})$ can be an SPT. MST is a special case of SPT, where $\vec{\beta}' = s\vec{\beta}$.
SPT covers a large range of transformation. However, not every function that satisfies the aforementioned conditions is physically meaningful: Consider the Jacobi matrix of transformed lens equation 1.14,

$$A'(\vec{\theta}) = \frac{\partial \vec{\beta}'}{\partial \vec{\beta}} \frac{\partial \vec{\beta}}{\partial \vec{\theta}} \equiv B(\vec{\beta}(\vec{\theta}))A(\vec{\theta}), \tag{1.15}$$

where $B(\vec{\beta})$ is the Jacobi matrix of the SPT. As A' is not guaranteed to be a symmetric matrix, it is also not guaranteed that there exists a mass distribution κ' which satisfies the deflection law after the SPT ($\kappa'_{,ij} \neq \kappa'_{,ji}$). One solution for the problem is to impose an axisymmetry of the lens to find SPTs that are exact: Furthermore, due to the uncertainties caused by incomplete information of local components, such as non-smooth components and inhomogeneous line-of-sight structure, various mass models that are not exactly in SPT relations can still reproduce the same lensing observables, to an observationally acceptable level. [104] has shown that an empirically known degeneracy between two widely used lens mass models, a power-law and a cored power-law, can be accounted for the SPT by showing that the differences between the lensing observables reproduced from these two models are not significant enough to be distinguished with an angular resolution of $\sim 5 \times 10^{-3}\theta_E$. Here, θ_E is the Einstein radius, and for galaxy-scale lensing, $\theta_E \sim 1''$.
However, [114] has shown with Illustris Simulation [108] that for the massive elliptical galaxies with velocity dispersion $\sigma > 250$ km/s, their density profile is close to a power law on average, with 10-20 % scatter, empirically justifying the use of "simple" lens mass model in the strong lensing regime.

Tables 1.1 and 1.2 summarize the type of degeneracies mentioned above and the changes in observables and models, respectively.

Other possible transformations are *Prismatic transformation* (also introduced in [39]), where both the source position and the deflection angles are changed by the same constant. Physically, it corresponds to having a very massive lens at a far distance transverse to the main deflector, and the source position is shifted to the same amount in the opposite direction to compensate for the change. However, this transformation is not considered as important as the situation is not very likely in practice.

Type of degeneracy	$\vec{\theta}$	μ_i/μ_j	Δt
ST (DD)	$\vec{\theta}$	invariant	$s\Delta t$
ST (AD)	$\sqrt{s}\vec{\theta}$	invariant	$s\Delta t$
MT (=MST)	$\vec{\theta}$	invariant	$s\Delta t$

Table 1.1: Summary of the degeneracies and their effect on observables. s is an arbitrary constant. PT is not included as it is not considered much in practice, and SPT is not listed as it is a family of transformations and cannot be summarized in a table.

Type of degeneracy	μ	$D_{\Delta t}$	κ **or** Σ	$\vec{\beta}$	L_s
ST (DD)	μ	s$D_{\Delta t}$	$\Sigma \rightarrow s\Sigma$	$\vec{\beta}$	$s^2 L_s$
ST (AD)	μ	$D_{\Delta t}$	$\Sigma \rightarrow s\Sigma$	$\sqrt{s}\vec{\beta}$	L_s
MT (=MST)	μ/s^2	$D_{\Delta t}$	$(1-\kappa) \rightarrow s(1-\kappa)$	$s\vec{\beta}$	L_s/s^2

Table 1.2: Summary of the mass model degeneracies and their effect on model quantities. L_s is the source luminosity.

1.2 Strong Lensing in Cosmology

1.2.1 Time-delay Distance

In 1964, Refsdal first proposed that strongly lensed type Ia supernovae can be used to determine cosmology, especially the Hubble constant, via the time-delay distance. This is a distance-like quantity that is a combination of three distances in a lens system [86]. The time-delay distance $D_{\Delta t}$ is defined as follows:

$$D_{\Delta t} \equiv \frac{D_{\rm d} D_{\rm s}}{D_{\rm ds}}. \tag{1.16}$$

Using this definition, equation 1.2 can be written as

$$\Delta t = \frac{1+z_{\rm L}}{c} D_{\Delta t} \phi(\vec{\theta}, \vec{\beta}). \tag{1.17}$$

Thus, when Δt and ϕ are constrained, the time-delay distance can also be constrained. Each of the three distances in equation 1.16 are inversely proportional to the Hubble constant, thus $D_{\Delta t}$ also follows the same inverse proportionality.

In calculating the time-delay distance, the main challenge is i) to model the mass distribution of the lens, or the Fermat potential $\phi(\vec{\theta}, \vec{\beta})$, from high-resolution imaging data, and ii) to obtain the time delay Δt from high-cadence monitoring. Having multiple extended images from an extended source, one can reconstruct the mass by imposing

that pixels from each image should be mapped to the source plane and reconstruct the same source. However, as stated in section 1.1.2, the mass model of the lens has various degeneracies, which are the main source of uncertainties in constraining cosmology using the time-delay distance [98]. Some of the mass model degeneracies are shown to be broken either by combining stellar kinematics data [55, 50], or by modeling the lens environment using cosmological simulations [88, 41].

The power of the time-delay distance in cosmology comes from the fact that it contains information at two different redshifts: the lens and the source. As it will be shown later in this thesis, in combination with the angular diameter distance, the time-delay distance can break the degeneracy between the dark energy equation of state and the curvature of the universe [51, 64]. As distance depends on cosmological models, the time-delay distance will allow us to constrain cosmological parameters, such as the Hubble constant in the low-redshift Universe.

Independent measurements of H_0 are receiving much attention of the cosmology community, as there are ongoing discussions on recently reported tensions between H_0 as measured from Planck and from local distance measurements. For example, [13] claimed that the tension can indicate modified early-time physics, or differences between the normalization of the two distance ladders anchoring the low redshift (H_0) and high redshift (sound horizon scale) Universe. Ongoing / completed analysis of five lensed quasars with well-measured time delays and mass models has been conducted by the H0 Lenses in COSMOGRAIL's Wellspring collaboration (H0LiCOW; [101]). The main goal of H0LiCOW is to measure the Hubble constant with $< 3.5\%$ precision using strong time-delay lenses, and to establish a baseline of strong lensing cosmography for upcoming observations that will increase the number of time-delay lenses significantly.

1.2.2 Probing Dark Matter Substructures

As stated in section 1.1.1, the relative magnification between images is an important observable in strong lensing. The deflection angle (which determines the image position) is the first derivative of the projected potential, while the magnification is proportional to the derivative of the image position, which is the second derivative of the potential. Thus magnification is more sensitive to local changes in the mass distribution than the image positions, while image positions are mainly determined by the global distribution of the smooth component. Anomalies in flux ratio have been used to probe clumpy and dark substructures in lens galaxy halos [66, 70]. It is difficult to measure the flux ratio anomalies of extended galaxies: thus, due to the compactness of the source, only strongly lensed quasars have been studied as systems that exhibit flux ratio anomalies. There have been discussions about the discrepancy between the substructure fraction from dark matter only simulations and the observed number of systems with flux ratio anomalies. For example, [113] have found that in comparison to what is reported in observations [25], the amount of flux ratio anomalies is underpredicted in two high-resolution cold dark matter simulations, Aquarius [95] and Phoenix [36]. Also, a recent work [42] has shown that some flux ratio anomalies can be explained by adding macro structure to the lens mass model, such as an

edge-on disk, which is motivated by deep observations using Keck-II Adaptive Optics and the Very Large Baseline Array (VLBA). This paper has proven that the observed flux ratio anomalies can argue for the existence of clumpy substructures only when more detailed observations of individual systems are available to rule out the possibility of complex macro models of the lens.

With deeper observations, fluctuations in surface brightness of the lensed images allow direct *gravitational imaging* [105, 56] that yields the mass and the position of the substructure. In this case, as the distortion is visible, the ambiguity between the clumpy substructure and the complex macro structure of the lens can be disentangled. For gravitational imaging technique, the mass detection threshold is set by the S/N and the resolution of the images as well as the position of the substructure with respect to the Einstein ring. The current detection limits on the masses of substructures in main halos are between 10^8 and $10^9 M_\odot$ depending on the position of the substructure, based on the Wide-Field Channel (WFC) of the Advanced Camera for Surveys (ACS) on the Hubble Space Telescope (HST). With the forthcoming Very Large Baseline Interferometer (VLBI) at radio wavelengths and the European Extremely Large Telescope (E-ELT) at optical wavelengths, the angular resolution and hence the detection of substructures via lenses will be improved significantly [107].

As different dark matter models predict different masses and abundances of substructures in a halo, detections and non-detections of substructures through flux ratio anomalies and gravitational imaging can be used to discriminate between dark matter models [106, 107]. The Strong lensing at High Angular Resolution Program (SHARP; [58]) has been obtaining high resolution imaging data of known strong lenses in the optical and near-infrared using HST and Keck adaptive optics for this purpose.

Chapter 2

Time-delay Lenses as Distance Indicators

In this chapter I present a novel method to infer cosmological parameters by using strong lenses as a standard ruler for measuring angular diameter distances to lenses. The idea was first proposed in [84] with a singular isothermal sphere for the lens mass model. I show how the method works analytically with an extended model. The contents of this chapter are published in [50].

2.1 Model

2.1.1 The idea: a simple analysis using singular isothermal spheres

We review the basic idea with the simplest case in which the mass density profile of a lens galaxy is given by an SIS. This case has been worked out by Paraficz and Hjorth in 2009 [84]. The density distribution of an SIS lens, $\rho_{\rm SIS}$, is given by

$$\rho_{\rm SIS}(r) = \frac{\sigma^2}{2\pi G r^2}, \tag{2.1}$$

where σ^2 is the three-dimensional isotropic velocity dispersion. The Einstein ring radius, θ_E, is related to σ^2 via

$$\sigma^2 = \theta_{\rm E}\frac{c^2}{4\pi}\frac{D_{\rm s}}{D_{\rm ds}}. \tag{2.2}$$

Clearly, the relation between the two observable quantities, $\theta_{\rm E}$ and σ, depends on the distance *ratio*.

To extract the actual angular diameter distance to the lens, $D_{\rm d}$, instead of the ratio, we need to include the lensing time delay [86]. The presence of intervening mass between the observer and the source, usually galaxies and/or clusters of galaxies, causes two different components on time delay: the *geometrical time delay* and the *potential time delay*. Strongly lensed systems show multiple images as photons coming from the source take

different paths: images are located at the closest approach to the lens of each path. The geometrical part of the time delay is caused by the fact that the total path lengths differ, while the potential part is caused by the difference in the depths of potential at each image position of the path.

In a SIS lens, the time delay between two images can be written as

$$\Delta t_{i,j} \equiv t_i - t_j = \frac{1+z_{\rm L}}{2c}\frac{D_{\rm d}D_{\rm s}}{D_{\rm ds}}(\theta_j^2 - \theta_i^2), \tag{2.3}$$

where θ_i is the angular separation between the i-th image and the center of the lens galaxy, and t_i is the absolute time delay of the i-th image, i.e., the delay in comparison to the case where the lens is absent [110]. The distance ratio that appears in this relation is the time-delay distance, $D_{\Delta t} \equiv (1+z_{\rm L})D_{\rm d}D_{\rm s}/D_{\rm ds}$, which depends primarily on H_0 and has a limited sensitivity to the other cosmological parameters, such as the equation of state of dark energy.

Remarkably, when we combine the above equation with equation (2.2) and $\theta_E = (\theta_i + \theta_j)/2$, we obtain the angular diameter distance to the lens:

$$D_{\rm d}(\theta_j - \theta_i) = \frac{c^3 \Delta t_{i,j}}{4\pi\sigma^2(1+z_{\rm L})}. \tag{2.4}$$

The physical interpretation of the above analysis is as follows: the velocity dispersion is determined by the gravitational potential of the lens, GM/r. The time delay gives the mass of the lens system, GM, and thus dividing them gives the physical size of the system, r. Since the angular scale of the system is directly observable via $\theta_j - \theta_i$, one can estimate the angular diameter distance to the lens. Equation (2.4) indeed gives the angular diameter distance as $D_{\rm d} \propto \Delta t_{i,j}/[\sigma^2(\theta_j - \theta_i)]$; thus, the uncertainty in $D_{\rm d}$ is given by the quadrature sum of the uncertainties in the time delay, velocity dispersion, and image position measurements.

As the velocity dispersion uncertainty is usually the biggest of all uncertainties, the uncertainty in $D_{\rm d}$ is expected to be dominated by the velocity dispersion uncertainty. The goal of this paper is to extend this analysis to more general lenses.

2.2 More realistic lenses

The analysis in section 2.1.1 assumes the simplest possible lens system: an SIS density profile with an isotropic velocity dispersion. While the SIS profile is widely used to model lens galaxies and is considered as a good approximation, several studies have shown that slopes of density profiles of individual galaxies show a non-negligible scatter from the SIS [54, 11, 12, 94]. In this section, we consider an arbitrary power-law density profile (section 2.2.1) to show that, in such a model, we can still extract $D_{\rm d}$ from $\Delta t_{i,j}$, σ^2, and image positions. We then show that the external convergence cancels out (section 2.2.2). We note that spherical symmetry is assumed throughout the paper.

2.2.1 Arbitrary slope of the spherical lens mass profile

Studies of early type galaxies (ETGs) as lenses have shown that the averaged *total* mass density profiles can be well approximated as a power-law, and also typical ellipticity of galaxies is fairly small [54, 11, 12, 94, 19]. Thus we allow the total mass density of a lens to follow a general power-law with spherical symmetry:

$$\rho = \rho_0 \Big(\frac{r}{r_0}\Big)^{-\gamma'}. \tag{2.5}$$

The distribution becomes a SIS for $\gamma' = 2$ (section 2.1.1). The lens potential also has a power-law form, $\psi(\theta) \propto \theta^l$, with $l = 3 - \gamma'$. The scaled deflection, $\vec{\alpha}$, which is given by $\nabla\psi = \vec{\alpha}$, and the lens equation, $\vec{\beta} = \vec{\theta} - \vec{\alpha}$, gives

$$\psi = \frac{1}{l}\,\vec{\theta}\cdot(\vec{\theta} - \vec{\beta}). \tag{2.6}$$

Using this result in equation (1.2), we obtain the time delay between two images as

$$\Delta t_{i,j} = \frac{1+z_{\rm L}}{2c}\frac{D_{\rm d}D_{\rm s}}{D_{\rm ds}}\left\{(\vec{\theta_i} - \vec{\beta})^2 - (\vec{\theta_j} - \vec{\beta})^2 - \frac{2}{l}\left[\vec{\theta_i}\cdot(\vec{\theta_i} - \vec{\beta}) - \vec{\theta_j}\cdot(\vec{\theta_j} - \vec{\beta})\right]\right\}. \tag{2.7}$$

From the geometry of the system, the lens equation and the definition of the angular diameter distance, the following relation between $\vec{\theta}$, $\vec{\beta}$, and $\hat{\alpha}$ holds:

$$\vec{\theta} - \vec{\beta} = \vec{\alpha} = \frac{D_{\rm ds}}{D_{\rm s}}\hat{\alpha}, \tag{2.8}$$

where $\hat{\alpha}$ is the deflection angle at the lens plane. We substitute $\vec{\theta} - \vec{\beta}$ in equation (2.7) for $\hat{\alpha}$, and write

$$\Delta t_{i,j} = D_{\rm d}\frac{(1+z_{\rm L})}{2c}\left[(\hat{\alpha}_i + \hat{\alpha}_j)\cdot(\vec{\theta_i} - \vec{\theta_j}) - \frac{2}{l}(\vec{\theta_i}\cdot\hat{\alpha}_i - \vec{\theta_j}\cdot\hat{\alpha}_j)\right]. \tag{2.9}$$

The remaining task is to relate $\hat{\alpha}$ to observables. As the potential of a spherically symmetric system only has a radial component with respect to the center, $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\theta}$ have only radial components. Let us define $\alpha \equiv |\hat{\alpha}|$, which is the magnitude of the deflection angle at the lens plane. Under the power-law density profile model, α is given by

$$\alpha(b) = \frac{2GM(b)}{c^2 b}F(\gamma') \propto b^{-\gamma'+2}, \tag{2.10}$$

where b is the physical separation between the lens and the point of the closest approach of the light ray, and

$$F(\gamma') \equiv \frac{\sqrt{\pi}\,\Gamma\left[\frac{1}{2}(-1+\gamma')\right]}{\Gamma(\frac{\gamma'}{2})}. \tag{2.11}$$

The derivation of this formula is given in appendix A.

Using the virial theorem, we obtain the radial velocity dispersion at a given radius r as

$$\sigma_r^2(r) = \frac{1}{2(\gamma' - 1)} \frac{GM(r)}{r} \propto r^{-\gamma'+2}. \tag{2.12}$$

If the velocity dispersion is isotropic, $\sigma_r^2(r) = \frac{1}{3}\sigma^2(r)$, and the radial velocity dispersion is the same as the line-of-sight velocity dispersion, which is observable. As both α and $\sigma_r^2(r)$ scale with radii in the same way, we can write $\alpha(b)$ as

$$\alpha = \frac{4(\gamma' - 1)}{c^2} F(\gamma') \sigma_r^2(b) = \frac{4(\gamma' - 1)}{c^2} F(\gamma') \sigma_r^2(r) \left(\frac{b}{r}\right)^{-\gamma'+2}. \tag{2.13}$$

We then obtain D_d from equation (2.9) with α given by equation (2.13),

$$D_\mathrm{d} = \frac{c^3 \Delta t_{i,j}}{4\pi \sigma_r^2(r)(1 + z_\mathrm{L})} (\Delta\tilde{\theta}_{i,j})^{-1}, \tag{2.14}$$

where[1]

$$(\Delta\tilde{\theta}_{i,j})^{-1} \equiv \frac{2\pi \left\{ \frac{2}{-\gamma'+3} \left[\theta_j \left(\frac{\theta_j}{\Theta}\right)^{-\gamma'+2} - \theta_i \left(\frac{\theta_i}{\Theta}\right)^{-\gamma'+2} \right] + (\theta_i + \theta_j) \left[\left(\frac{\theta_i}{\Theta}\right)^{-\gamma'+2} - \left(\frac{\theta_j}{\Theta}\right)^{-\gamma'+2} \right] \right\}^{-1}}{F(\gamma')(\gamma' - 1)}, \tag{2.15}$$

and Θ is the angular position at which the velocity dispersion is measured, i.e., $r = \Theta D_\mathrm{d}$. For $\gamma' = 2$, we obtain $\Delta\tilde{\theta}_{i,j} = \theta_j - \theta_i$, and thus we can reproduce the result of the SIS model (equation (2.4)).

Equation (2.14) still supports the basic physical picture that the ratio of $\Delta t_{i,j}$ and σ_r^2 gives some effective physical size of the lens, and dividing it by the appropriate angular separation in the sky, $\Delta\tilde{\theta}_{i,j}$, gives the angular diameter distance. The main difference between the SIS and the power law density profiles is that, in the latter case, the velocity dispersion is a function of radii. In general, image positions are different from the points at which the velocity dispersion is measured. Thus, we need to correct for the mismatch of the exact locations of the velocity dispersion measurement and the image positions. This is why the $\left(\frac{\theta}{\Theta}\right)^{-\gamma'+2}$ term appears in the final expression of D_d: it scales the velocity dispersion such that we can get the potential at the image position. This requires us to measure (or model) the density slope, γ', as well.

2.2.2 External convergence

In modeling realistic lens systems, one important factor to consider is the so-called "mass-sheet transformation (MST)". MST is a subset of the source-position transformation [92]. Degeneracy exists, such that there are many mass models of the lens galaxy that can

[1] We use $\hat{\theta}_i \cdot \hat{\theta}_j = -1$ in reducing the vector dot products in equation (2.9) to the scalar products in equation (2.15).

simultaneously reproduce most of the lensing observables, such as image positions and flux ratios, with different source positions [31]. This degeneracy constitutes one of the dominant sources of uncertainty in measuring the time-delay distance [98, 97, 100, 91]. In this subsection, we show that the effect of MST cancels out, leaving no effect on the inferred $D_{\rm d}$.

Once we choose a model for the convergence field, $\kappa_{\rm model}(\vec{\theta})$, that matches the observations, we transform $\kappa_{\rm model}$ and $\vec{\alpha}$ to obtain a new convergence field, $\kappa_{\rm MST}(\vec{\theta})$, and a new scaled deflection, $\vec{\alpha}_{\rm MST}$, as

$$\kappa_{\rm MST}(\vec{\theta}) = \lambda + (1-\lambda)\kappa_{\rm model}(\vec{\theta}), \tag{2.16}$$

$$\vec{\alpha}_{\rm MST}(\vec{\theta}) = \lambda\vec{\theta} + (1-\lambda)\vec{\alpha}_{\rm model}(\vec{\theta}) \tag{2.17}$$

$$= \lambda\vec{\theta} + \vec{\alpha}_{\rm MST,lens}(\vec{\theta}), \tag{2.18}$$

where λ is a constant which physically corresponds to the scaled convergence of a uniform sheet of mass external to the lens galaxy. In equation (2.18), we decompose the transformed deflection into two parts; a deflection from the lens, and that from the external convergence. We define $\vec{\alpha}_{\rm MST,lens} \equiv (1-\lambda)\vec{\alpha}_{\rm model}$, whose meaning will be explained later in this subsection. To satisfy the lens equation (1.8) while leaving the image positions invariant, the source position must transform as

$$\vec{\beta}_{\rm MST} = (1-\lambda)\vec{\beta}_{\rm model}, \tag{2.19}$$

which is why this transformation is a part of the family of transformation called the *source-position* transformation.

Considering the following relation among κ , ϕ and ψ,

$$\phi = \frac{1}{2}(\vec{\theta}-\vec{\beta})^2 - \psi, \tag{2.20}$$

$$\nabla^2\psi = 2\kappa, \tag{2.21}$$

the transformed Fermat potential of the i-th image, $\phi_{{\rm MST},i}$, becomes

$$\phi_{{\rm MST},i} = (1-\lambda)\phi_{{\rm model},i} - \frac{\lambda(1-\lambda)}{2}|\vec{\beta}|^2. \tag{2.22}$$

Since the source position $\vec{\beta}$ is the same for all the images, the second term in equation (2.22) cancels out if we calculate the difference in the Fermat potential between two images i and j. Thus, the difference, $\Delta\phi_{i,j}$, transforms as

$$\Delta\phi_{{\rm MST},i,j} = (1-\lambda)\Delta\phi_{{\rm model},i,j}. \tag{2.23}$$

As the time delay is directly proportional to the Fermat potential, we find that $\Delta t_{i,j}$ is simply increased by a factor of $1-\lambda$ after the MST for fixed distances/cosmology.

If we assume that the physical origin of MST is an effective external convergence due to mass structures along the line of sight, $\kappa_{\rm ext}$, we can identify λ with $\kappa_{\rm ext}$. In the following, we apply the MST to the power-law mass model and show that the inferred $D_{\rm d}$ remains unaffected by $\kappa_{\rm ext}$. We start first with the special case of SIS to gain intuition before considering the general power-law profile.

Singular isothermal sphere

Here we follow the steps from section 2.1.1, but with MST applied to it. From equation (2.16), the transformed density profile of the lens is

$$\rho_{\mathrm{SIS,MST}} = (1-\kappa_{\mathrm{ext}})\rho_{\mathrm{SIS,model}}. \tag{2.24}$$

Note that the original transformation equation (2.16) is written in terms of the convergence, κ; however, as the convergence and the density profile are proportional to each other (equation (1.5)), we transform the density in the same way as the convergence. To satisfy equation (2.1), the velocity dispersion must transform as

$$\sigma^2_{\mathrm{MST}} = (1-\kappa_{\mathrm{ext}})\sigma^2. \tag{2.25}$$

Equation (2.2) then becomes

$$\sigma^2_{\mathrm{MST}} = (1-\kappa_{\mathrm{ext}})\theta_{\mathrm{E}}\frac{c^2}{4\pi}\frac{D_{\mathrm{s}}}{D_{\mathrm{ds}}}. \tag{2.26}$$

From equation (2.23), the time-delay equation (2.3) transforms as

$$\Delta t_{\mathrm{MST},i,j} = (1-\kappa_{\mathrm{ext}})\frac{1+z_{\mathrm{L}}}{2c}\frac{D_{\mathrm{d}}D_{\mathrm{s}}}{D_{\mathrm{ds}}}(\theta_j^2-\theta_i^2), \tag{2.27}$$

and by combining the above two equations, we get

$$\Delta t_{\mathrm{MST},i,j} = \frac{4\pi}{c^3}\sigma^2_{\mathrm{MST}}(1+z_{\mathrm{L}})D_{\mathrm{d}}(\theta_i-\theta_j), \tag{2.28}$$

in which κ_{ext} cancels out. This equation is identical to equation (2.4), but with the transformed quantities, $\Delta t_{\mathrm{MST},i,j}$ and σ^2_{MST}.

The reason is as follows. Suppose that we have a lens system which has a velocity dispersion of σ^2 and the time-delay difference of Δt. We then try to model this system by a lens plus an external convergence, κ_{ext}. Then, the modeled σ and Δt would be different from the original ones by a factor of $1-\kappa_{\mathrm{ext}}$, but the ratio of the two is invariant. As D_{d} is proportional to the ratio of the two, we can measure the same D_{d} as before, regardless of the existence of the external convergence.

Spherical power-law density profile

Now we study the effect of MST on the spherical power-law density profile lens galaxy model, following section 2.2.1. Combining the time-delay transformation with equation (2.9) yields

$$\begin{aligned}\Delta t_{\mathrm{MST},i,j} =& (1-\kappa_{\mathrm{ext}})\Delta t_{\mathrm{model},i,j}\\ =& (1-\kappa_{\mathrm{ext}})D_{\mathrm{d}}\frac{(1+z_{\mathrm{L}})}{2c}\\ &\times\left[(\hat{\alpha}_{\mathrm{model},i}+\hat{\alpha}_{\mathrm{model},j})\cdot(\vec{\theta}_i-\vec{\theta}_j)-\frac{2}{l}(\vec{\theta}_i\cdot\hat{\alpha}_{\mathrm{model},i}-\vec{\theta}_j\cdot\hat{\alpha}_{\mathrm{model},j})\right].\end{aligned} \tag{2.29}$$

Again, the density normalization of the lens galaxy, ρ_0, transforms as

$$\rho_{0,\mathrm{MST}} = (1-\kappa_{\mathrm{ext}})\rho_{0,\mathrm{model}}, \tag{2.30}$$

and thus among the total deflection angle α, only a $(1-\kappa_{\mathrm{ext}})$ fraction of it is from the lens, which is why we denoted this contribution as $\vec{\alpha}_{\mathrm{MST,lens}} = (1-\lambda)\vec{\alpha}_{\mathrm{model}}$ in equation (2.18). Using this in equation (2.29) yields

$$\begin{aligned}\Delta t_{\mathrm{MST},i,j} &= D_{\mathrm{d}}\frac{(1+z_L)}{2c} \\ &\times\left[(\hat{\alpha}_{\mathrm{MST,lens},i}+\hat{\alpha}_{\mathrm{MST,lens},j})\cdot(\vec{\theta}_i-\vec{\theta}_j)-\frac{2}{l}(\vec{\theta}_i\cdot\hat{\alpha}_{\mathrm{MST,lens},i}-\vec{\theta}_j\cdot\hat{\alpha}_{\mathrm{MST,lens},j})\right].\end{aligned} \tag{2.31}$$

As the measured velocity dispersion of the lens gives the estimate of the lens potential only, the relation between the deflection angle from the lens and the velocity dispersion does not change after the MST:

$$|\hat{\alpha}_{\mathrm{MST,lens}}| = \frac{4(\gamma'-1)}{c^2}\sigma_r^2(R)F(\gamma')\left(\frac{b}{R}\right)^{-\gamma'+2}. \tag{2.32}$$

Thus, D_{d} can be calculated from the original equation (2.14) even after the MST.

This is an important finding. In the previous studies of the time-delay distance to measure the Hubble constant, κ_{ext} was the main obstacle in measuring H_0 precisely [98]. On the other hand, we have shown that D_{d} measured from strong lensing, which combines the time-delay, the image position, and the velocity dispersion data, does not suffer from the effect of κ_{ext}.

2.3 Error formula and implications for B1608+686 and RXJ1131−1231

2.3.1 Aperture-averaged line of sight velocity dispersion

We do not measure the radial component of the velocity dispersion, $\sigma_r^2(r)$. Rather, we measure the luminosity-weighted line-of-sight velocity dispersion, $\sigma_p^2(R)$. We relate them using the following equation:

$$\sigma_p^2(R) \equiv I_p(R)\sigma_s^2(R) = 2\int_R^\infty\left(1-\beta_{\mathrm{ani}}\frac{R^2}{r^2}\right)\frac{\rho_*(r)\sigma_r^2(r)rdr}{\sqrt{r^2-R^2}}. \tag{2.33}$$

Here, r denotes the three-dimensional radius, while R denotes the projected radius. We shall use these two different radii notations for the rest of the paper. β_{ani} is the effect of the velocity dispersion anisotropy, which will be studied in detail in section 2.4. In this section, we set $\beta_{\mathrm{ani}} = 0$. The other functions are: $I_p(R)$ is the projected stellar distribution function, $\sigma_s(R)$ is the projected velocity dispersion and $\sigma_r(r)$ is given by equation (2.12).

For a stellar density profile, ρ_*, we consider two profiles that are known to describe well the stellar light distributions of galaxies: the Hernquist profile and the Jaffe profile. These two different profiles would also allow us to assess the effect of luminosity weighting on $\sigma_{\rm p}^2(R)$.

A generalized form of the stellar density distribution, which satisfies $\rho_* \propto r^{-4}$ as $r \to \infty$, can be expressed as

$$\rho_* = \frac{(3-\gamma_s)I_0}{4\pi}\frac{a}{r^{\gamma_s}(r+a)^{4-\gamma_s}}, \tag{2.34}$$

where $0 \leq \gamma_s < 3$, following [26]. The Hernquist profile corresponds to $\gamma_s = 1$ [40]:

$$\rho_*(r) = \frac{I_0 a}{2\pi r(r+a)^3}, \tag{2.35}$$

where I_0 is a normalization and a is a scale radius determined by $a = (2^{1/(3-\gamma_s)} - 1)r_{1/2}$ following [26], where $r_{1/2}$ is the half-mass radius. Due to the projection effect, the two-dimensional half-light radius $R_{\rm eff}$ is related to the three-dimensional half-mass radius $r_{1/2}$ as $r_{1/2} = 1.33R_{\rm eff}$ for the Hernquist profile, thus the scale radius $a = 0.551R_{\rm eff}$. The projected Hernquist distribution is known to provide a good fit for the stellar distribution of elliptical galaxies that follow the de Vaucouleurs law,

$$I_p(R) = \frac{I_0}{2\pi a^2(1-s^2)^2}[(2+s^2)X(s) - 3], \tag{2.36}$$

where $s \equiv R/a$ is a scaled projected radius, and $X(s)$ is defined as

$$X(s) \equiv \begin{cases} \dfrac{1}{\sqrt{1-s^2}}\,\mathrm{sech}^{-1}\,s & \text{for } 0 \leq s \leq 1 \\ \dfrac{1}{\sqrt{s^2-1}}\,\sec^{-1}\,s & \text{for } 1 \leq s < \infty \end{cases} . \tag{2.37}$$

We then examine the case where the stellar density profile has a steeper slope at the center, by using the Jaffe model [49], which has $\gamma_s = 2$. We do not consider models with $2 < \gamma_s < 3$ as they fail to represent the basic physical properties of a galaxy, e.g. diverging potential / velocity dispersion at the center. In the Jaffe model, the stellar density profile becomes

$$\rho_* = \frac{1}{4\pi}\frac{a}{r^2(r+a)^2}, \tag{2.38}$$

and the projected surface brightness distribution, $I_p(R)$, becomes

$$I_p(R) = \frac{I_0}{4a^2 s} - \frac{I_0}{2\pi a^2}\frac{1}{s^2-1}\left[(s^2-2)X(s)+1\right], \tag{2.39}$$

where for the Jaffe profile $r_{1/2} = 1.31R_{\rm eff}$ and $a = 1.31R_{\rm eff}$.

We note that both the Hernquist and Jaffe profiles for the stars are not single power-laws, and neither are dark matter distributions such as the Navarro, Frenk and White

profile [77]. Stars and dark matter have different radial distributions in galaxies with the stars typically dominating over dark matter at the central parts and vice versa at outer parts. The contributions of stars and dark matter are often comparable around the effective radius. Despite the different radial distributions of stellar and dark matter, the total density profile of stars and dark matter is remarkably well described by a power-law within a few effective radius, as previous lensing and/or dynamical studies have shown (e.g.,[54, 12, 100, 19]). Therefore, our use of the Hernquist/Jaffe profiles for the luminosity weighting to scale the velocity dispersion measured near the effective radius to the image positions is consistent with the use of a power-law for the total density profile. For a total density profile that is nearly isothermal, there is an inconsistency in the slope between the total density profile and the Hernquist profile at the center ($r \lesssim 0.1$). However, the contribution of the enclosed mass from this central region to the total enclosed mass within either the Einstein radius or effective radius (approximately where we have lensing/dynamical measurements) is insignificant. Thus, the central slope inconsistency between the stellar and the total density has negligible impact on our lensing and dynamical analysis.

Ideally, we wish to measure the line-of-sight velocity dispersion profile as a function of projected radii. In practice, however, most of the observations do not allow us to spatially resolve the galaxy; rather they allow us to measure the luminosity-weighted, aperture-averaged velocity dispersion inside an aperture of a fixed size [18]. We calculate the luminosity-weighted aperture-averaged projected velocity dispersion, $\langle\sigma_p^2\rangle_{\rm ap}$, as follows:

$$\langle\sigma_p^2\rangle_{\rm ap} \equiv \frac{\int_{\rm ap} I_p \sigma_s^2 R\, dR\, d\theta}{\int_{\rm ap} I_p R\, dR\, d\theta}. \tag{2.40}$$

2.3.2 Analytic formula

In this section, we relate the statistical uncertainty in $D_{\rm d}$ to those of the observables, i.e., Δt, σ_p^2, and γ'. (The effect of an anisotropic velocity dispersion will be discussed in detail in section 2.4.) Assuming that these observables are independently measured, we write the total uncertainty in $D_{\rm d}$, hereafter $S_{D_{\rm d}}$, as

$$\begin{aligned} S_{D_{\rm d}} =& \sqrt{\left(\frac{\partial D_{\rm d}}{\partial \Delta t_{i,j}}\right)^2 S_{\Delta t}^2 + \left(\frac{\partial D_{\rm d}}{\partial \sigma_p^2}\right)^2 S_{\sigma_p^2}^2 + \left(\frac{\partial D_{\rm d}}{\partial \gamma'}\right)^2 S_{\gamma'}^2} \\ =& D_{\rm d}\sqrt{\left(\frac{1}{\Delta t_{i,j}}\right)^2 S_{\Delta t}^2 + \left(\frac{1}{\sigma_p^2}\right)^2 S_{\sigma_p^2}^2 + \frac{1}{D_{\rm d}^2}\left(\frac{\partial D_{\rm d}}{\partial \gamma'}\right)^2 S_{\gamma'}^2}, \end{aligned} \tag{2.41}$$

where S_x is the measurement uncertainty in the variable x. Since image positions, $\theta_{i,j}$, are precisely measured, we do not include their uncertainties in this formula. In the following sections 2.3.3 and 2.3.4, we shall apply this formula to two lens systems, B1608+656 and RXJ1131−1231, respectively.

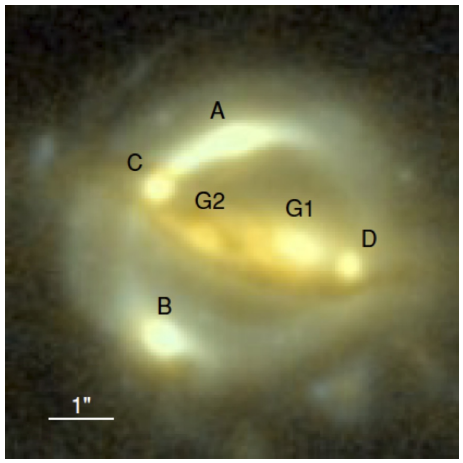

Figure 2.1: Image of B1608+656, adopted from figure 1 of [98].

2.3.3 B1608+656

Figure 2.1 shows the image configuration of B1608+656 [99]. The information on image configuration is important as our formula applies only to a circularly symmetric case. Thus, the only image pairs we can use in this paper are the ones that are on the opposite sides of the lens center. More thorough analysis using all the data will be presented elsewhere (Suyu et al., in preparation). The data of B1608+656 are mostly from [98], but the image positions are calculated from the data given in [55], the time delays are from [32], and the redshifts are from refs. [33, 74]. For this system, the origin of the coordinates is set at the image A. The data are summarized as :

$$\begin{aligned}
z_{\rm L} =& 0.6304 \\
z_{\rm s} =& 1.394 \\
\vec{\theta}_{\rm A} =& (0.0'', 0.0'') \\
\vec{\theta}_{\rm B} =& (-0.7380'', -1.9612'') \\
\vec{\theta}_{\rm C} =& (-0.7446'', -0.4537'') \\
\vec{\theta}_{\rm D} =& (1.1284'', -1.2565'') \\
R_{\rm eff} =& 0.58'' \\
\gamma' =& 2.08 \pm 0.03 \\
\langle \sigma_p^2 \rangle_{\rm ap}^{1/2} =& 260 \pm 15 \text{ km/s} \\
\Delta t_{\rm AB} =& 31.5^{+2.0}_{-1.0} \text{ days} \\
\Delta t_{\rm CB} =& 36.0^{+1.5}_{-1.5} \text{ days} \\
\Delta t_{\rm DB} =& 77.0^{+2.0}_{-1.0} \text{ days} \\
\Delta t_{\rm CD} =& \Delta t_{\rm CB} - \Delta t_{\rm DB} = -41.0^{+2.5}_{-1.8} \text{ days}.
\end{aligned} \tag{2.42}$$

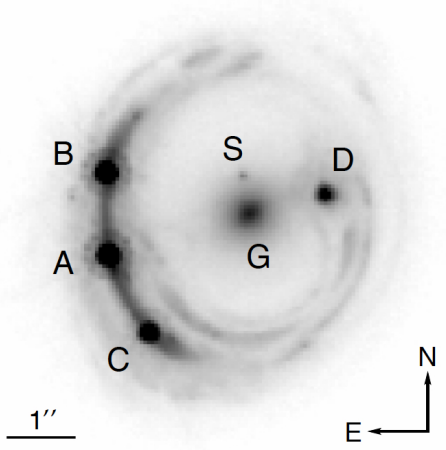

Figure 2.2: Image of RXJ1131−1231, adopted from figure 1 of [97].

We use the CD pair. Also, as we write $D_{\rm d}$ in terms of $\sigma_r(r)$ (e.g. 2.14), we normalize the radial velocity dispersion profile, $\sigma_r(r)$, using $\langle\sigma_p^2(R)\rangle_{\rm ap}$ given by the observation and using equations (2.33) and (2.40). With these values, we find $D_{\rm d} = 1485.7$ Mpc. For comparison, $D_{\rm d}$ from the best-fit WMAP 7-year parameters is $D_{\rm d} = 1423.3$ Mpc. We now use equation (2.41) to compute $S_{D_{\rm d}}$:

$$S_{D_{\rm d}} = D_{\rm d}\sqrt{3.72\times10^{-3} + 1.33\times10^{-2} + 2.36\times10^{-3}}, \tag{2.43}$$

where from the first term, each number indicates the fractional uncertainty in $D_{\rm d}$ contributed by the time-delay measurement $\Delta t_{i,j}$, the line-of-sight velocity dispersion measurement σ_p^2, and the density profile index γ'. (Note that $S_{\sigma_p^2}/\sigma_p^2 = 2S_{\sigma_p}/\sigma_p$.) With this value, the total uncertainty, including all the terms in equation (2.43), is $S_{D_{\rm d}} = 0.14D_{\rm d}$, i.e., 14% uncertainty. The dominant contribution comes from the uncertainty in σ_p, which gives $S_{D_{\rm d}} = 0.12D_{\rm d}$.

2.3.4 RXJ1131−1231

In this section we repeat the same analysis as above, but with another well-studied strong lensing time-delay system, RXJ1131−1231, using data from refs. [102, 93] for the time delays and the redshifts, respectively, and from [97] for the other quantities. The data for

this system are

$$\begin{aligned}
z_{\rm L} =& 0.295 \\
z_{\rm s} =& 0.658 \\
\vec{\theta}_{\rm G} =& (4.411'', 4.011'') \\
\vec{\theta}_{\rm A} =& (2.383'', 3.412'') \\
\vec{\theta}_{\rm D} =& (5.494'', 4.288'') \\
R_{\rm eff} =& 1.85'' \\
\gamma' =& 1.95^{+0.05}_{-0.04} \\
\langle \sigma_p^2 \rangle_{\rm ap}^{1/2} =& 323 \pm 20 \text{ km/s} \\
\Delta t_{\rm AB} =& 0.7 \pm 1.4 \text{ days} \\
\Delta t_{\rm DB} =& 91.4 \pm 1.5 \text{ days} \\
\Delta t_{\rm AD} =& \Delta t_{\rm AB} - \Delta t_{\rm DB} = -90.7 \pm 2.1 \text{ days}.
\end{aligned} \tag{2.44}$$

We use the AD pair. Using these values, we find $D_{\rm d} = 813.33$ Mpc, and $D_{\rm d}$ from the best-fit WMAP 7-year parameters is $D_{\rm d} = 876.5$ Mpc. The total uncertainty in $D_{\rm d}$ is

$$S_{D_{\rm d}} = D_{\rm d}\sqrt{5.36 \times 10^{-4} + 1.53 \times 10^{-2} + 1.46 \times 10^{-3}} = 0.13 D_{\rm d}. \tag{2.45}$$

The velocity dispersion alone gives $S_{D_{\rm d}} = 0.12 D_{\rm d}$.

Therefore, we expect the *existing* data on these systems to yield $D_{\rm d}$ with $13 - 14\%$ precision per object, assuming the isotropic velocity dispersion. In the next section, we shall study the effect of the largest source of systematic uncertainty in our method: an anisotropic velocity dispersion, and how to reduce its effect in the estimation of $D_{\rm d}$.

2.4 Anisotropic velocity dispersion

The anisotropic stellar motion changes the relation between the potential and the observed line-of-sight velocity dispersion. As our method crucially relies upon knowing the potential depth, we must take into account the anisotropic velocity dispersion of stars. We do this by following [98], which uses spherical Jeans modeling to relate the observed line-of-sight velocity dispersion to the mass distribution. We then study the effect of anisotropy on the aperture-averaged value of the velocity dispersion (section 2.4.1) as well as on the velocity dispersion measured at the so-called "sweet spot" (section 2.4.2). Finally, we use Monte Carlo simulations to compute the effect of anisotropy on the uncertainty in $D_{\rm d}$ (section 2.4.3).

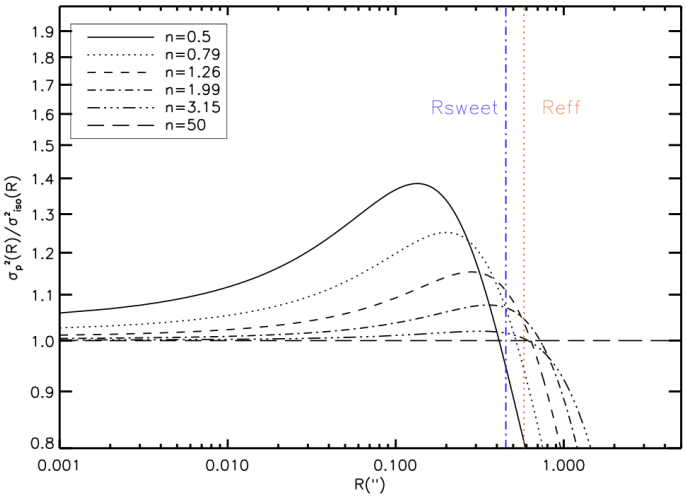

Figure 2.3: Ratio of $\sigma_p^2(R)$ to $\sigma_{\rm iso}^2(R)$, as a function of the projected radius R, and $n \equiv r_{\rm ani}/R_{\rm eff}$. The former is observable, while the latter is related more directly to GM/R. Two vertical lines show the effective radius ($R_{\rm eff}$) and the sweet-spot radius ($R_{\rm sweet}$) defined in section 2.4.2.

2.4.1 Spherical Jeans equation

We solve the spherical Jeans equation for a given mass distribution (i.e., a power-law density profile) to obtain the three-dimensional radial velocity dispersion σ_r,

$$\frac{1}{\rho_*}\frac{d(\sigma_r^2\rho_*)}{dr} + 2\beta_{\rm ani}\frac{\sigma_r^2}{r} = -\frac{GM(\leq r)}{r^2}. \tag{2.46}$$

Here, the anisotropy function, $\beta_{\rm ani}(r)$, is chosen as the Osipkov-Merritt (OM) anisotropy [83, 68],

$$\beta_{\rm ani}(r) \equiv \frac{r^2}{r_{\rm ani}^2 + r^2} = 1 - \frac{\sigma_T^2(r)}{\sigma_r^2(r)}, \tag{2.47}$$

where $\sigma_T(r)$ and $\sigma_r(r)$ are the velocity dispersions in the tangential and radial directions, respectively. Although the anisotropy is parametrized by a single variable, $r_{\rm ani}$, under this specific model, we can model almost any velocity structures by linearly superimposing the solutions [68]. We then calculate $\sigma_p^2(R)$ from $\sigma_r^2(r)$ using equation 2.33, and $\langle\sigma_p^2\rangle_{\rm ap}$ using equation 2.40.

To quantitatively demonstrate the behavior of the anisotropic velocity dispersion, we again use the observations of B1608+656 introduced in section 2.3.3 for the analysis in this and the following sections.

In figure 2.3, we show the ratio of $\sigma_p^2(R)$ to the isotropic velocity dispersion, $\sigma_{\rm iso}^2(R)$, with $a = 0.551R_{\rm eff}$ and $R_{\rm eff} = 0.58''$ for the Hernquist profile. The isotropic velocity dispersion is a solution to the Jeans equation (2.46) with no anisotropy, $\beta_{\rm ani} \equiv 0$; thus, it

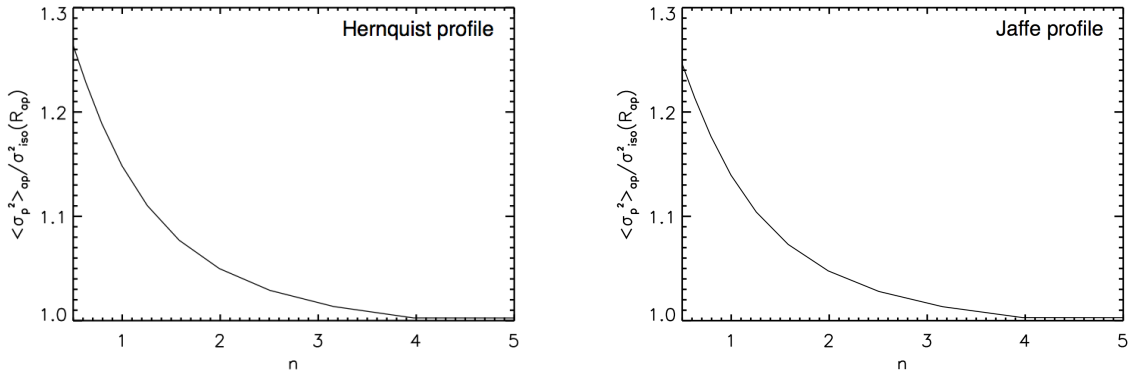

Figure 2.4: Ratio of $\langle\sigma_p^2\rangle_{\rm ap}$ to $\sigma^2_{\rm iso}(R_{\rm ap})$, as a function of n. For the stellar density distribution, we use the Hernquist and Jaffe profiles in the left and right panels, respectively. The size of the aperture is fixed at $R_{\rm ap} = 0.42''$. The upper limit of n, 5, is chosen since the velocity dispersion does not differ much from the isotropic case beyond n of 5, while the lower limit, 0.5, is determined by observations (e.g. [57]) and radial instability arguments (e.g. [69, 96]). The unknown anisotropy dominates the uncertainty on $\langle\sigma^2_{\rm p}\rangle_{\rm ap}$.

is related more directly to GM/R. We have one free parameter, n, which parametrizes the anisotropic radius as

$$r_{\rm ani} \equiv nR_{\rm eff}. \tag{2.48}$$

For a given mass distribution of the lens, $\sigma_p^2(R)$ depends on n. We vary n from 0.5 to 50 in logarithmic spacing. We find $\sigma_p^2(R)/\sigma^2_{\rm iso}(R) \approx 1$ to within 10% at $R = R_{\rm eff}$, except for the highly anisotropic case of $n = 0.5$ when the stellar distribution follows Hernquist profile.

In figure 2.4, we show the ratio of $\langle\sigma_p^2\rangle_{\rm ap}$ to $\sigma^2_{\rm iso}(R_{\rm ap})$ as a function of n, where $R_{\rm ap}$ is fixed to 0.42". In the left panel, this ratio reaches 26% for $n = 0.5$, and decreases as n increases when the stellar distribution follows the Hernquist profile. In the right panel, we show the same ratio for the Jaffe stellar distribution, with the ratio reaching 24% for $n = 0.5$. Overall the difference in $\langle\sigma_p^2\rangle_{\rm ap}$ between Hernquist and Jaffe distributions is small compared to the impact of the anisotropy. Therefore, for the remainder of the paper, we consider only the Hernquist distribution as a conservative model, where the dominant uncertainty on $\langle\sigma_p^2\rangle_{\rm ap}$ is due to the unknown anisotropy.

Since the inferred $D_{\rm d}$ is proportional to the inverse of the isotropic velocity dispersion, having a large variation in the inferred isotropic velocity dispersion can cause a large uncertainty in $D_{\rm d}$. Unfortunately, anisotropy is not directly observable, unless we have a three-dimensional velocity dispersion measurement. Clearly, a better approach is needed.

2.4.2 Sweet-spot method

It has been pointed out that, when the observations of the surface brightness profile and the velocity dispersion profile are available, one can find the so-called *sweet spot*, $R_{\rm sweet}$,

at which the effect of the anisotropic velocity dispersion on the mass determination is minimized [23]. Also see [111, 109]. The OM anisotropy model has an isotropic core and a radial envelope. However, as we observe the projected velocity dispersion, there are two components that play roles in the estimation of the observed velocity dispersion. The anisotropy changes the ratio between tangential and radial components of the velocity dispersion at a given radius, while the projection changes the magnitude of contributions from radial and tangential components. Quantitatively, at a fixed radius of observation R, $\sigma_p(R)$ has contributions from infinitely many shells with radii $r = R/\cos x$, where $x = [0, \pi/2]$. At each radius r we can decompose the contributions to the projected velocity dispersion into tangential and radial components as $\sigma_T(r)\cos x$ and $\sigma_r(r)\sin x$, respectively. Due to the weighting by the trigonometric functions, at small x, contributions from the tangential component is bigger than that from the radial component, and vice versa at large x. Now, let us assume that the total velocity dispersion, $\sigma^2(r) = \sigma_T^2(r) + \sigma_r^2(r)$, is the same for the isotropic and anisotropic model for a given galaxy mass, as it is proportional to the total kinetic energy. Then, when $r = R/\cos x$ is large, the tangential component is suppressed while the radial component is enhanced compare to the isotropic case, due to the anisotropy (since $\sigma_T(r)$ becomes small for large r in equation (2.47)). As a result, in comparison to the isotropic case where $\sigma_T(r) = \sigma_r(r)$, anisotropic velocity dispersion shows $\sigma_p(R) > \sigma_{\rm iso}(R)$ at small R, and $\sigma_p(R) < \sigma_{\rm iso}(R)$ at large R. Thus, if we observe an anisotropic system, there exists a projected radius R at which the transition from one to the other occurs, as we increase R from the center of a galaxy to the outskirt of it. This transition radius is the sweet spot.

While the analytical derivation of $R_{\rm sweet}$ has been done assuming a constant $\beta_{\rm ani}$, the further study [65] shows that the method works for systems with a non-constant $\beta_{\rm ani}$ as well. The sweet spot can be determined from the brightness profile of a massive elliptical galaxy [65]. It is close to the projected radius at which R satisfies $d\ln I(R)/d\ln R = -2$. For a Hernquist surface brightness profile, we find $R_{\rm sweet} \approx 0.78R_{\rm eff}$. It is also shown in [23] that while the Sersic index changes from 1 to 12, $R_{\rm sweet}$ varies only about $0.3R_{\rm eff}$, thus the sweet-spot radius is fairly insensitive to the luminosity profile. In figure 2.3, the sweet-spot radius is shown as the left vertical line. We find that the difference between projected velocity dispersions with various anisotropy parameters is minimum around $R = 0.78R_{\rm eff} = 0.45''$ with the data of B1608+656. It particularly reduces the effect of a highly anisotropic case with $n = 0.5$, compared to using $\sigma_p^2(R)$ at the effective radius or the aperture-averaged σ_p^2. The uncertainty in the mass of massive ellipticals estimated from the sweet-spot method is claimed to be 5-7 per cent. Therefore, the best approach is to use spatially-resolved spectroscopic data of lens galaxies to obtain the velocity dispersion at the sweet spot.

2.4.3 Monte Carlo simulation

We use Monte Carlo simulations to study how much the velocity anisotropy inflates the uncertainty in $D_{\rm d}$, and how well we can mitigate it by using the sweet-spot method.

For two time-delay systems B1608+656 and RXJ1131−1231, we generate 11 discrete radial profiles of anisotropic velocity dispersion by solving the Jeans equation, with log-

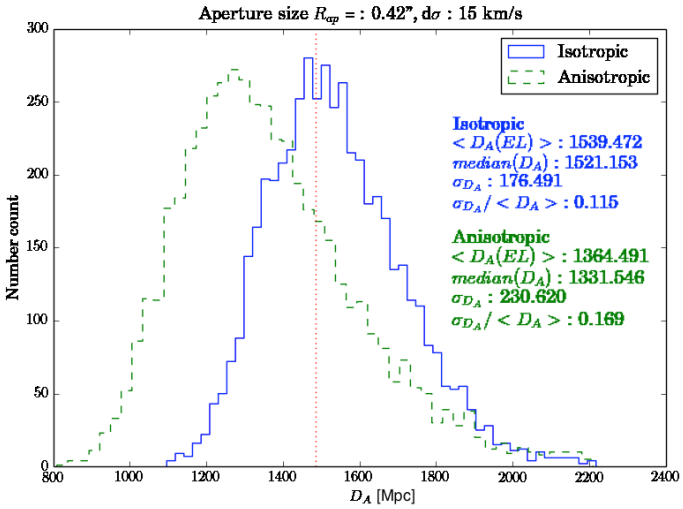

Figure 2.5: Simulated distribution of $D_{\rm d}$ to B1608+656. The solid and dashed histograms show the distributions with the isotropic and anisotropic simulations, interpreted by the isotropic model. We use the aperture averaged velocity dispersion, $\langle\sigma_p^2\rangle_{\rm ap}$, with the aperture size of $0.42''$. The standard deviation of the velocity dispersion used in simulations is 15 km/s. The fractional uncertainty in $D_{\rm A}$ is 11.5% in the case of isotropic velocity dispersion model, while in the case of anisotropic velocity dispersion model it is 16.9%.

arithmically spaced $n = [0.5, 50]$. The effective radius and the density profile index γ' are fixed at the best-fit values given in sections 2.3.3 and 2.3.4. We randomly choose one profile from the set of different anisotropy parameters to create a mock galaxy. We then compute σ_p^2 from each mock galaxy in three ways: the aperture-averaged value $\langle\sigma_p^2\rangle_{\rm ap}$ with the aperture size of $0.42''$ for both systems, $\sigma_p^2(R)$ with $R = R_{\rm eff}$, and $\sigma_p^2(R)$ with $R = R_{\rm sweet}$. As the uncertainty in $D_{\rm d}$ is dominated by that of σ_p^2, we add a Gaussian random noise to σ_p^2 with variance of $S^2_{\sigma_p^2} = 2S^2_{\sigma_p}(S^2_{\sigma_p} + 2\sigma_p^2)$. We then compute $D_{\rm d}$ from these simulated data with the best-fit values of the time-delay data and image positions given in sections 2.3.3 and 2.3.4. (We do not add noise to time delays or image positions.) While our simulated galaxies have anisotropic velocity dispersions, we use the isotropic velocity dispersion model to calculate $D_{\rm d}$. In this way we can quantify the effect of our ignoring anisotropic velocity dispersion by marginalizing over it.

Figures 2.5, 2.6, and 2.7 show the distributions of $D_{\rm d}$ obtained from mock B1608+656 realizations using $\langle\sigma_p^2\rangle_{\rm ap}$, $\sigma_p^2(R_{\rm eff})$, and $\sigma_p^2(R_{\rm sweet})$, respectively. The solid and dashed histograms in each panel show the realizations with the isotropic and anisotropic velocity dispersions, respectively. The former realizations are used to check validity of our simulations, as well as to make a direct assessment of the effect of anisotropy. The vertical dotted lines show $D_{\rm d} = 1485.7$ Mpc that we obtained in section 2.3.3.

We summarize the results from the analysis on B1608+656 and RXJ1131−1231 in tables 2.1 and 2.2, respectively. The uncertainties in $D_{\rm d}$ from isotropic simulations (interpreted

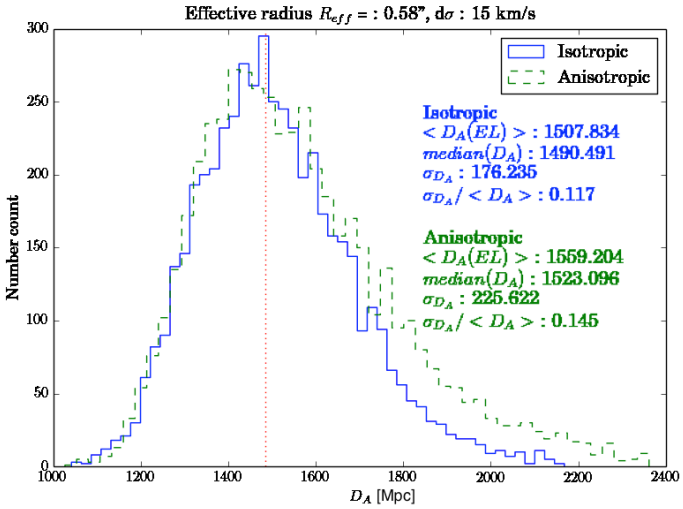

Figure 2.6: Same as figure 2.5, but with $\sigma_p^2(R_{\rm eff})$ and $R_{\rm eff} = 0.58''$. The fractional uncertainty in $D_{\rm A}$ is 11.7% in the case of isotropic velocity dispersion model, while in the case of anisotropic velocity dispersion model it is 14.5%.

by the isotropic model) agree with the analytical estimates given in sections 2.3.3 and 2.3.4. On the other hand, those from anisotropic simulations (again interpreted by the isotropic model) show significantly larger uncertainties when $\langle\sigma_p^2\rangle_{\rm ap}$ or $\sigma_p^2(R_{\rm eff})$ is used. Fortunately, using $\sigma_p^2(R_{\rm sweet})$ eliminates most of the inflation of the uncertainty due to velocity anisotropy.

Figure 2.7 shows that the peak is shifted in the anisotropic case in comparison to the isotropic case, while in figure 2.6 the peak remains at the same position. This is due to the marginalization of the anisotropy. In figure 2.3, we choose 6 different n values that are spaced logarithmically, and choose two radii ($R_{\rm eff}$ and $R_{\rm sweet}$) to calculate the $D_{\rm d}$ distributions for both the isotropic and anisotropic cases. At $R_{\rm sweet}$, the scatter between the curves is smaller compare to that at $R_{\rm eff}$; however, at $R_{\rm sweet}$, the curves are also shifted toward higher velocity dispersions compared to the isotropic case. As a result, the whole distribution of $D_{\rm d}$ is shifted toward lower values. On the other hand, at $R_{\rm eff}$, while the scatter is larger, there is no systematic change in σ_p^2 relative to $\sigma_{\rm iso}$ (i.e. among 6 values of n, two give σ_p^2 larger than the $\sigma_{\rm iso}^2$, two give smaller, and the other two give σ_p^2 almost identical to the $\sigma_{\rm iso}^2$ value). As a result, the peak position does not change, while we get an extended tail towards higher $D_{\rm d}$ value. This does not mean that using $R_{\rm sweet}$ gives a biased $D_{\rm d}$, as we cannot assume that the velocity dispersion structure is isotropic. Also, as the width of the distribution is much bigger than the shift of the peak, at the moment the effect of this shift is negligible. As the distribution of $D_{\rm d}$ depends on the choice of the anisotropy model as well as on the range/selection of n, we study another anisotropy parameterization to see the robustness of the results against the choice of parameterization in the next section.

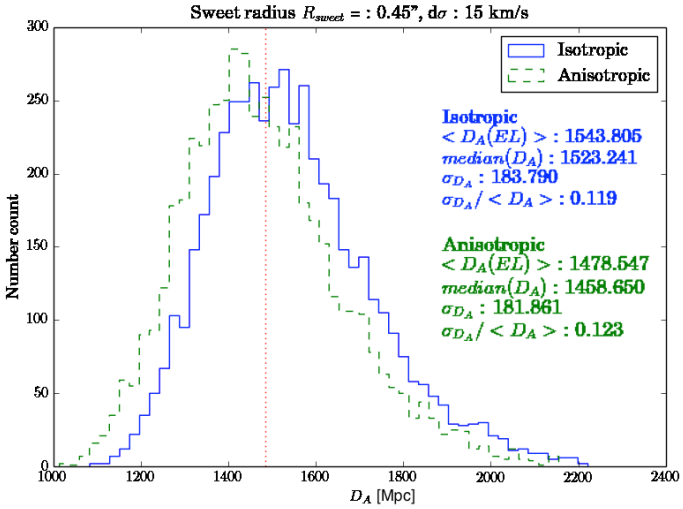

Figure 2.7: Same as figure 2.5, but with $\sigma_p^2(R_{\rm sweet})$ and $R_{\rm sweet} = 0.45''$. The fractional uncertainty in $D_{\rm A}$ is 11.9% in the case of isotropic velocity dispersion model, while in the case of anisotropic velocity dispersion model it is 12.3%.

Table 2.1: Expected fractional uncertainty in $D_{\rm d}$ to B1608+656

	Isotropic	**Anisotropic**
$R_{\rm ap}$	11.5%	16.9%
$R_{\rm eff}$	11.7%	14.5%
$R_{\rm sweet}$	11.9%	12.3%

2.4.4 Two Parameter Extension Model

To show that the sweet spot is not a unique characteristic of OM anisotropy, we repeat the same analysis using a different spatially-varying anisotropy parameter, $\beta_{\rm ani}(r)$, from [5] :

$$\beta_{\rm ani}(r) = \frac{\beta_{\rm in} r^2 + \beta_{\rm out} r_{\rm a}^2}{r^2 + r_{\rm a}^2}, \tag{2.49}$$

which we will call the TPE (Two Parameter Extension) model. Two additional parameters, $\beta_{\rm in}$ and $\beta_{\rm out}$, are added to the OM anisotropy. The OM model takes $\beta_{\rm in} = 0$ and $\beta_{\rm out} = 1$, so as $r \to 0$ the velocity dispersion becomes isotropic, while at the outskirt the velocity dispersion becomes radial. We see that $\beta_{\rm ani} \to \beta_{\rm in}$ toward the center, and $\beta_{\rm ani} \to \beta_{\rm out}$ toward the outskirts of the galaxy. We follow [5] and adopt flat priors on $\beta_{\rm in} = [-0.6, 0.6]$ and $\beta_{\rm out} = [-0.6, 0.6]$, while the anisotropic radius, $r_{\rm a}$, is scaled in the same way as in the OM model (equation 2.48).

The resulting velocity dispersion profiles are shown in figure 2.8. Near the sweet spot, the fractional uncertainty in the velocity dispersion becomes as small as 15%. Also we note that the deviation from the isotropic velocity dispersion is not skewed at the sweet

Table 2.2: Expected fractional uncertainty in $D_{\rm d}$ to RXJ1131−1231

	Isotropic	**Anisotropic**
$R_{\rm ap}$	12.5%	15.1%
$R_{\rm eff}$	12.4%	14.8%
$R_{\rm sweet}$	12.5%	12.6%

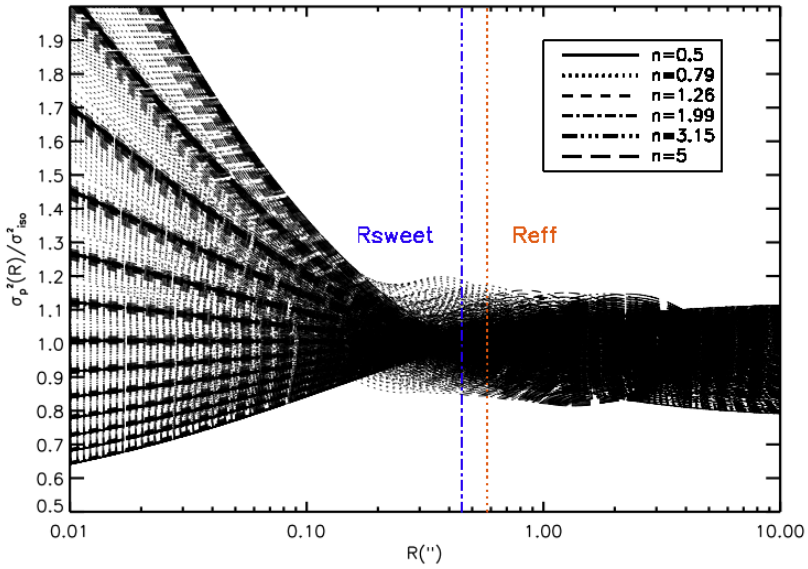

Figure 2.8: Same as figure 2.3, but with TPE anisotropy model. The range of the two new parameters, $\beta_{\rm in}$ and $\beta_{\rm out}$, is $[-0.6, 0.6]$ for both parameters, with steps of $\delta\beta_{\rm in} = 0.1$ and $\delta\beta_{\rm out} = 0.1$.

spot, which keeps the peak of the posterior distribution of $D_{\rm d}$ at the same place as for the isotropic dispersion model. The posterior distribution calculated at the sweet-spot radius is shown in figure 2.9. We find that the uncertainty on angular diameter distance using this parametrization is about 13% for B1608+656, and 14% for RXJ1131−1231, comparable to those in section 2.4.3.

2.5 Conclusion

We have shown that we can determine $D_{\rm d}$ to strong lens systems with time delays. The underlying physics is simple; thus, this method offers a robust determination of $D_{\rm d}$ to individual systems. The key advantage of this method is that the external convergence does not affect the distance determination. The uncertainty in the inferred $D_{\rm d}$ is dominated by that in the velocity dispersion and its anisotropy. The effect of anisotropy can be minimized

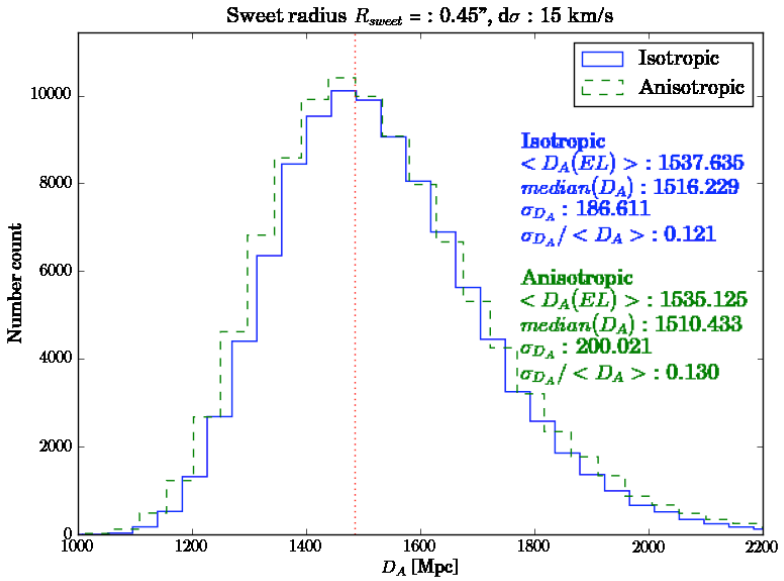

Figure 2.9: Same as figure 2.7, but with TPE anisotropy model. Note that we use more realizations here in comparison to the previous analysis, as the parameter combination is 169 times as much as the one from OM parameterization, due to two additional parameters. As a consequence, the result for the isotropic case is slightly different from figure 2.7. The fractional uncertainty in $D_{\rm A}$ is 12.1% in the case of isotropic velocity dispersion model, while in the case of anisotropic velocity dispersion model it is 13.0%.

by measuring the velocity dispersion at the sweet-spot radius.

The *existing* data on B1608+656 and RXJ1131−1231 should yield $D_{\rm d}$ with 17% and 15% precision, respectively. If we use the velocity dispersions at the sweet-spot radii, the precision improves to about 13%. In figure 2.10, we show the expected fractional uncertainty in $D_{\rm d}$ to B1608+656 as a function of the uncertainty in the velocity dispersions, σ. The σ at the sweet-spot radius measured with 260 ± 7 km/s corresponds to σ^2 measured with 5% precision. This yields $D_{\rm d}$ with 7% precision, after marginalizing over velocity anisotropy. We show the robustness of our results using two different parameterizations of $\beta_{\rm ani}(R)$, but a further study may be needed for more general cases.

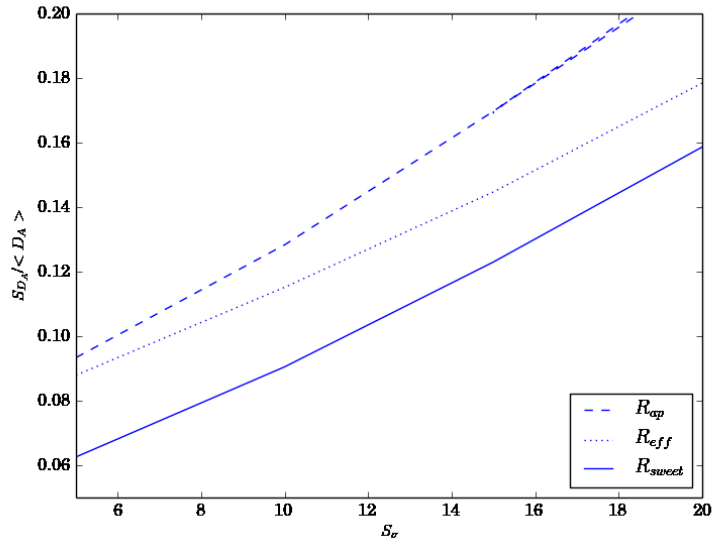

Figure 2.10: Expected fractional uncertainty in $D_{\rm d}$ to B1608+656 as a function of the uncertainty in σ in units of km/s. The dashed, dotted, and solid lines are for $\sigma = \langle\sigma_p^2\rangle_{\rm ap}^{1/2}$, $\sigma_p(R_{\rm eff})$, and $\sigma_p(R_{\rm sweet})$, respectively.

Chapter 3

Implementations and Measurements

The method presented in the previous chapter has been implemented in the full analysis pipeline used by refs. [98, 97]. We show how dynamical models have been incorporated into lensing chains, and present the results in this chapter. The contents are based on Jee et al. in preparation.

3.1 Importance Sampling the Lensing Chain with Dynamical Models

To constrain the angular diameter distance using observational data, we use the posterior probability density function (PDF) for the lens model parameters obtained via Monte Carlo Markov Chain (MCMC) analysis using the lensing and time-delay data in [98, 100] for B1608+656 and RXJ1131−1231, respectively. For the kinematics of the lens, we importance sample [61, 98] the posterior PDF for the lens mass model parameters with the likelihood of the kinematics data using Jeans modeling with anisotropic velocity dispersion and with flat priors on the anisotropy parameters, as shown in section 2.4. The model velocity dispersion is luminosity weighted and aperture averaged to be compared with the observation. In this section I briefly summarize the data, models, and the marginalization process, following the notation in [100].

Time-delay lensing cosmography requires multiple types of observations per lens, each providing different information about the lens. First, the imaging data provide the surface brightness distribution of the lens and the lensed source in a pixelated form, which enables the reconstruction of the Fermat potential via modeling of the lens mass distribution. For both lenses, HST Advanced Camera for Surveys (ACS) data were used. The imaging data is denoted as $\boldsymbol{d}_{\mathbf{ACS}}$. This data include the lens light and the lensed source, which is separated into extended host galaxy surface brightness and the point-like AGN light for each AGN image. These components are blurred by the point spread function, which can be modeled using stars in the field of view. Long-term monitoring of the variability of individual images provides the time-delay between image pairs, $\boldsymbol{\Delta t}$. For B1608+656, the time delay is measured by [33, 32]. For RXJ1131−1231, a new time-delay measurement

by monitoring in optical wavelengths is provided in [102]. The spectroscopic data for the lenses are obtained from the Low-Resolution Imaging Spectrometer (LRIS; [82]) on Keck 1. The spectroscopic data provide the luminosity-weighted, aperture-averaged line-of-sight velocity dispersion of the lens, $\langle\sigma_{ap}^2\rangle^{1/2}$. For convenience, we denote this simply as σ in this section. The process of obtaining the velocity dispersion from spectroscopic data is described in [100]. In measuring the time-delay distance, the lens environment, $\boldsymbol{d}_{\mathbf{env}}$, should be modeled to break the mass-sheet degeneracy [31, 98, 91, 100]. This provides information on the external convergence $\kappa_{\rm ext}$.

The lens mass distribution is modeled as an elliptical power-law profile with external shear, the lens light as an elliptical Sersic profile, and the lensed source light is modeled with an independent point-source AGN per image plus an extended source surface brightness on a regular grid for the host galaxy. Once the lens mass distribution is constrained from the image, the time delay between image pairs given the lens mass model can be calculated. Thus, time delay pairs provide some constraints on the lens model as well as constraining the time-delay distance. More details on the lens mass model and time delay are provided in [98, 100]. We model the velocity dispersion using two parametric anisotropy models as introduced in the previous chapter. The external convergence is estimated based on the relative number count of galaxies neighboring the lens with respect to the field galaxies, and calibrated using the Millennium Simulation [41].

Under the Bayesian framework, with the choice of physically motivated models for each component, the problem then becomes finding the posterior PDF of the model parameter set $\boldsymbol{\xi}$,

$$P(\boldsymbol{\xi}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}). \tag{3.1}$$

The model parameters are $\boldsymbol{\xi} = \{\boldsymbol{D}, \gamma', \theta_E, \gamma_{ext}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\rm ext}\}$, where $\boldsymbol{D} = \{D_d, D_{ds}/D_s\}$, and $\boldsymbol{\eta}$ is a vector containing the rest of the lens model parameters that do not affect kinematics. Two differences between the parameters we use here and the previous analysis in [98, 100] are: i) Instead of calculating the posterior PDF of cosmological parameters (denoted as $\boldsymbol{\pi}$ in the previous paper), we chose to directly calculate the posterior PDF of D_d and D_{ds}/D_s. This is because we want to constrain the angular diameter distance alone, while the previous analysis focused on getting cosmological parameters and calculating the joint constraint of the lensing time-delay distances and the Cosmic Microwave Background (CMB) obtained from the Wilkinson Microwave Anisotropy Probe (WMAP) data. ii) We test two velocity dispersion anisotropy models, thus instead of using $r_{\rm ani}$, we use $\boldsymbol{\beta}_{\mathbf{ani}}$.

Bayes' theorem states

$$P(\boldsymbol{\xi}|\boldsymbol{d}) = \frac{P(\boldsymbol{\xi})\ P(\boldsymbol{d}|\boldsymbol{\xi})}{E(\boldsymbol{d})} \tag{3.2}$$

where $\boldsymbol{\xi}$ is the parameter vector, $\boldsymbol{d}$ is the data vector, $P(\boldsymbol{\xi})$ is the prior, $P(\boldsymbol{d}|\boldsymbol{\xi})$ is the likelihood of the data $\boldsymbol{d}$ given the parameters and $E(\boldsymbol{d})$ is the model evidence, which can be used to compare different models. By applying Bayes' theorem to equation 3.1, the proportionality between the posterior PDF and likelihood can be stated as follows:

$$P(\boldsymbol{\xi}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}) \propto P(\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}|\boldsymbol{\xi})P(\boldsymbol{\xi}). \tag{3.3}$$

As all observations are independent, the likelihood in equation 3.3 can be separated as

$$P(\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}|\boldsymbol{\xi}) = P(\boldsymbol{d}_{\mathbf{ACS}}|\boldsymbol{\xi})P(\sigma|\boldsymbol{\xi})P(\boldsymbol{d}_{\mathbf{env}}|\boldsymbol{\xi})P(\boldsymbol{\Delta t}|\boldsymbol{\xi}). \tag{3.4}$$

Considering the data dependency of the parameters, and again, using Bayes' theorem, the above equation can be written further as

$$\begin{aligned} P(\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}|\boldsymbol{\xi}) \propto P(\boldsymbol{d}_{\mathbf{ACS}}, \boldsymbol{\Delta t}|\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \kappa_{\text{ext}})P(\sigma|\boldsymbol{D}, \gamma', \theta_E, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}) \\ P(\boldsymbol{d}_{\mathbf{env}}|\gamma_{\text{ext}}, \kappa_{\text{ext}})P(\boldsymbol{D})P(\gamma')P(\theta_E)P(\gamma_{\text{ext}})P(\boldsymbol{\eta})P(\boldsymbol{\beta}_{\mathbf{ani}})P(\kappa_{\text{ext}}). \end{aligned} \tag{3.5}$$

To obtain the posterior PDF of $\boldsymbol{D}$, we marginalize the full posterior $P(\boldsymbol{\xi}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t})$ over the other parameters,

$$P(\boldsymbol{D}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}) = \int d\gamma'\ d\theta_E\ d\gamma_{\text{ext}}\ d\boldsymbol{\eta}\ d\boldsymbol{\beta}_{\mathbf{ani}}\ d\kappa_{\text{ext}} P(\boldsymbol{\xi}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}), \tag{3.6}$$

by *importance sampling*. Importance sampling states that the expectation value of a function, $f(x)$ where x follows the PDF $P_1(x)$ can be calculated in the following way:

$$\langle f(x)\rangle_{P_1} = \int P_1(x)f(x)dx = \int \frac{P_1(x)}{P_2(x)}P_2(x)f(x)dx = \langle \frac{P_1(x)}{P_2(x)}f(x)\rangle_{P_2}, \tag{3.7}$$

where only $P_2(x)$ is available and $P_1(x)$ is not. In combination with Bayes' theorem, importance sampling allows us to separate the data in the posterior distribution. For example, in our case, $P_1 = P(\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t})$, which is the posterior from all the available data and $P_2 = P(\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}|\boldsymbol{d}_{\mathbf{ACS}}, \boldsymbol{\Delta t})$, which is the posterior only from the lens image and time-delay data. Bayes' theorem yields

$$\begin{aligned} P(\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}) \\ \propto P(\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}|\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}) \end{aligned} \tag{3.8}$$

and due to the independence of the individual observation as in equation 3.4, the likelihood can be separated as follows:

$$\begin{aligned} P(\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}|\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}) =& P(\boldsymbol{d}_{\mathbf{ACS}}, \boldsymbol{\Delta t}|\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}) \\ & P(\sigma|\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}})P(\kappa_{\text{ext}}|\boldsymbol{d}_{\mathbf{env}}). \end{aligned} \tag{3.9}$$

Applying Bayes' theorem again, we obtain

$$\begin{aligned} P(\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}|\boldsymbol{d}_{\mathbf{ACS}}, \sigma, \boldsymbol{d}_{\mathbf{env}}, \boldsymbol{\Delta t}) \propto& P(\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}}|\boldsymbol{d}_{\mathbf{ACS}}, \boldsymbol{\Delta t}) \\ & P(\sigma|\boldsymbol{D}, \gamma', \theta_E, \gamma_{\text{ext}}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\text{ext}})P(\kappa_{\text{ext}}|\boldsymbol{d}_{\mathbf{env}}), \end{aligned} \tag{3.10}$$

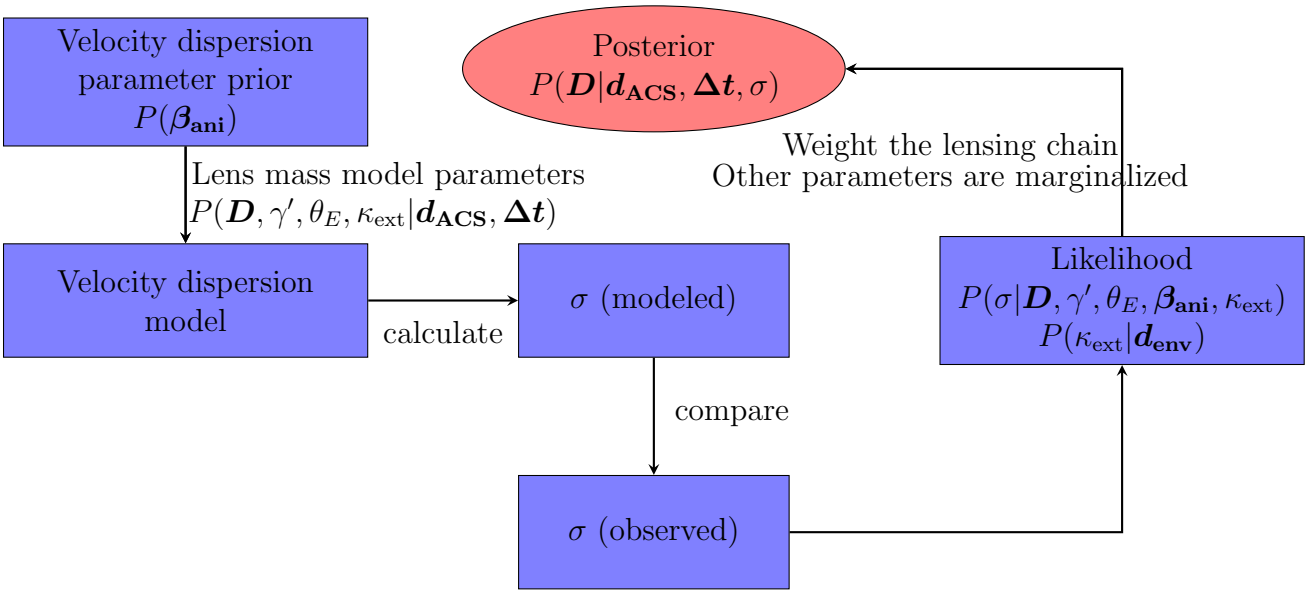

Figure 3.1: Schematic diagram of importance sampling, showing that the dynamical model is separately calculated and incorporated into the lensing chain, to constrain the posterior PDF of the distances.

and thus

$$\frac{P_1}{P_2} \propto P(\sigma|\boldsymbol{D}, \gamma', \theta_E, \gamma_{\rm ext}, \boldsymbol{\eta}, \boldsymbol{\beta}_{\mathbf{ani}}, \kappa_{\rm ext}) P(\kappa_{\rm ext}|\boldsymbol{d}_{\mathbf{env}}). \tag{3.11}$$

This shows that when calculating the posterior distribution of $\boldsymbol{D}$, the lensing time-delay chain can be weighted by the likelihood of σ and $\boldsymbol{d}_{\mathbf{env}}$, then integrated over the rest of parameters. For more details, see e.g. [61, 98].

As shown above, the kinematics likelihood can be calculated separately and incorporated into the lensing chain. The results we present in section 3.2 compare OM and TPE velocity dispersion models for each lens.

We give flat priors to $D_{\rm d}$, $D_{\rm ds}/D_{\rm s}$ and used PDF of $\kappa_{\rm ext}$ from [41] as shown in figure 3.6. Figure 3.1 schematically summarizes the process.

3.2 Measurements

We show the result of our analysis to constrain the angular diameter distance with current data. Our main results show i) the constraints on angular diameter distance are made by including the kinematic data, and ii) the independence of angular diameter distance to the external convergence (figures 3.2 - 3.5). We find that for B1608, the 1σ uncertainty on measured angular diameter distance was similar for both models: for OM profile it is 13%, and for TPE model it is 14%. However, for RXJ1131-1231, it is 14% for OM profile while the uncertainty increases to 18% as we use the TPE model, which we further discuss in section 3.2.2.

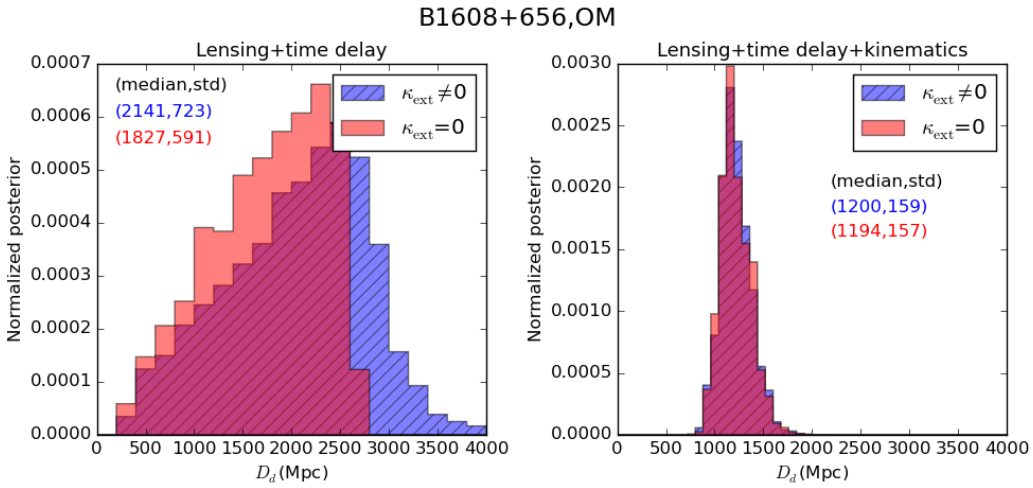

Figure 3.2: The posterior distribution of the angular diameter distance to the lens B1608+656. The left panel only includes lensing and time-delay information, while the right panel shows the result combined with kinematics of the lens, which gives 13% measurement of the angular diameter distance to the lens. The blue hatched distribution is when the external convergence distribution is estimated by ray-tracing through the Millennium Simulation [41], while the red is when the external convergence is set to be zero. The similarity between red and blue distributions in the right panel shows that by combining the kinematic information with lensing time delays, we can measure the angular diameter distance that is essentially independent of $\kappa_{\rm ext}$.

3.2.1 Independence of the external convergence

We show in figures 3.2 - 3.5 the angular diameter distance constrained using the method, without (left panel) and with (right panel) the velocity dispersion information. While the external convergence shifts the posterior distribution of the distance without the velocity dispersion, the distance becomes independent of the external convergence when the velocity dispersion information is used. This is due to the fixation of the gravitational potential from the velocity dispersion information, which normalizes the angular diameter distance. The distances are constrained to 13-18%, for two systems with two anisotropy models. The estimate on the distribution of external convergence for each lens is provided following [41], which uses nearby galaxy number counts around the lens, which are then calibrated by ray-tracing through the similar fields in the Millennium Simulation. We show the constraints on the Hubble constant in figure 3.10.

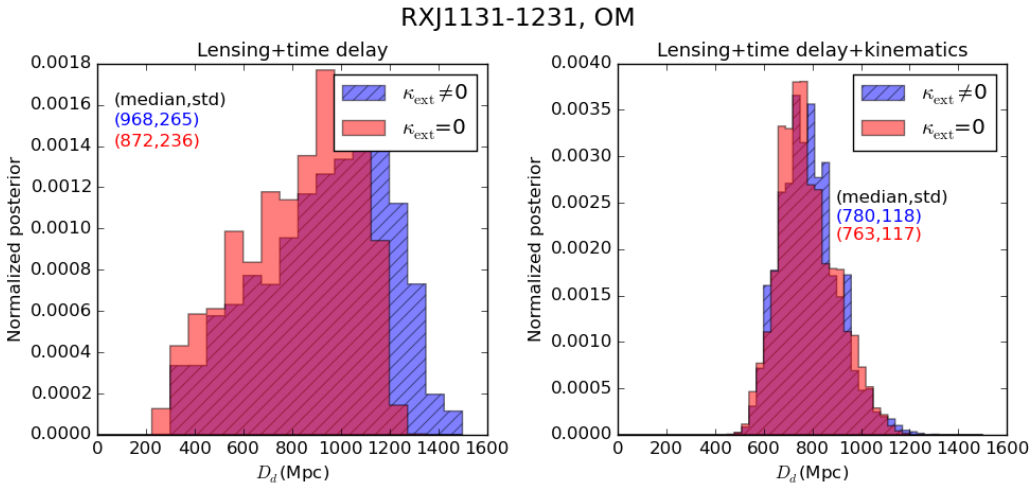

Figure 3.3: The same as figure 3.2, but for RXJ1131-1231. The right panel gives 15% measurement of the angular diameter distance to the lens.

3.2.2 Optimal aperture size to effective radius ratio

Current observations of the lens galaxies provide the velocity dispersion measured over an aperture of fixed physical size. The velocity dispersion is estimated via spectroscopy on RXJ1131-1231 with a rectangular aperture of width 0.81″, where the center of the aperture is placed at the center of the lens galaxy. For RXJ1131-1231, the effective radius of the galaxy is $r_{\rm eff} = 1.85''$, thus the width of the aperture is ∼22% of the effective radius. For B1608+656, the width of the aperture is ∼72% of the effective radius. In Figure 3.7, we vary the width of the aperture to show how anisotropic models of velocity dispersion changes the predicted aperture-averaged velocity dispersion. As this is to test the effect of the aperture size on the precision of D_d, we keep the rest of the parameters fixed. With an infinite aperture width, the observed velocity dispersion reaches the virial limit, where the uncertainty due to the anisotropy will be minimized and the difference due to the density profile is the only factor determining the aperture averaged velocity dispersion [5]. However, the physical width of the aperture should be limited due to spatial broadening. The slit disperses incoming photons along the direction of the width of the slit, thus having a wide slit causes mixing of spatial and spectral dimension, which effectively causes extra broadening. Spatial broadening thus increases the measurement uncertainty of velocity dispersion.

However, for the velocity dispersion data we have, where the size of the aperture is some fraction of the effective radius, the uncertainty due to the anisotropy is more significant. For example, the vertical dot-dashed line is where the $R_{\rm ap}/r_{\rm eff}$ corresponds to that of RXJ1131-1231 data. The red shaded region shows larger uncertainty than the blue one, which is why the posterior distribution of the angular diameter distance of RXJ1131-1231

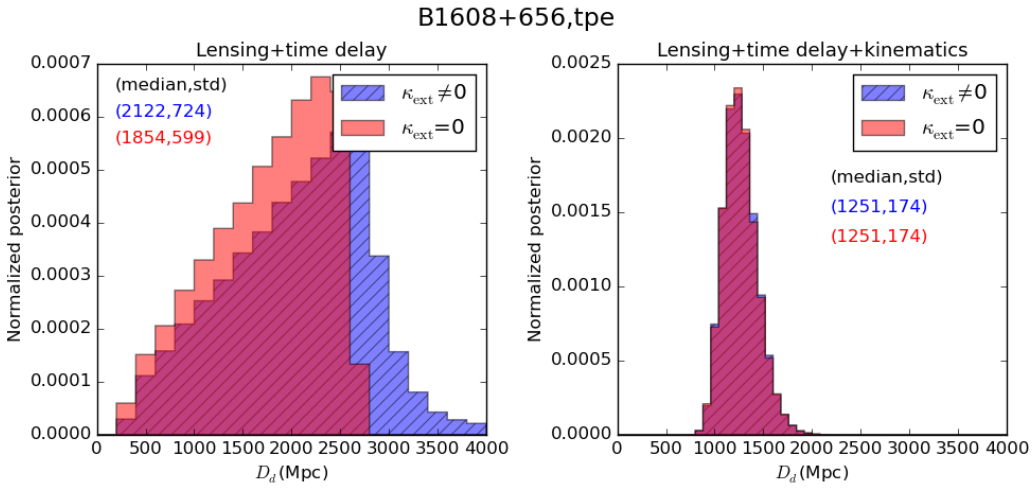

Figure 3.4: The same as figure 3.2, but for TPE anisotropy model. The right panel gives 14% measurement of the angular diameter distance to the lens.

with TPE model is wider than that of the OM model. The dotted line is when $R_{\rm ap}/r_{\rm eff}$ corresponds to B1608+656 value, where the width of two distributions are similar. In terms of anisotropy, we predict that the $R_{\rm ap}/r_{\rm eff} > 0.7$ is required to avoid uncertainty due to anisotropy being bigger than the measurement uncertainty. Current measurements were made with the aperture of half width $\sim 0.4''$, and with 20km/s measurement uncertainty in velocity dispersion, which is mostly from the systematic uncertainties in the measurement. This puts a limit on the effective radius of the lens galaxy that we can measure and model reliably with the current method: $r_{\rm eff} < 0.57''$. We note that however, depending on the anisotropy model, the mean of the velocity dispersion varies even with a large aperture to effective radius ratio, which can introduce bias in the posterior distribution of the angular diameter distance. With the current measurement uncertainty of velocity dispersion this is not a significant problem, and this indicates that we need to measure spatially resolved velocity dispersion to constrain both the velocity dispersion anisotropy and the mass profile simultaneously.

3.2.3 Cosmological implications : Constraining Hubble constant from inverse-distance ladder

The SNe type Ia provides an accurate shape of the expansion history of the Universe, as they yield distances at multiple redshifts. While it can only constrain relative distances between the SNe due to the systematic uncertainties, specifically in their intrinsic brightness, in combination with absolute distances at any redshift they span, the SNe can be normalized to yield absolute distances. Also, the absolute distance anchors the relative distances SNe measure, that can be propagated to the zero redshift to yield the Hubble

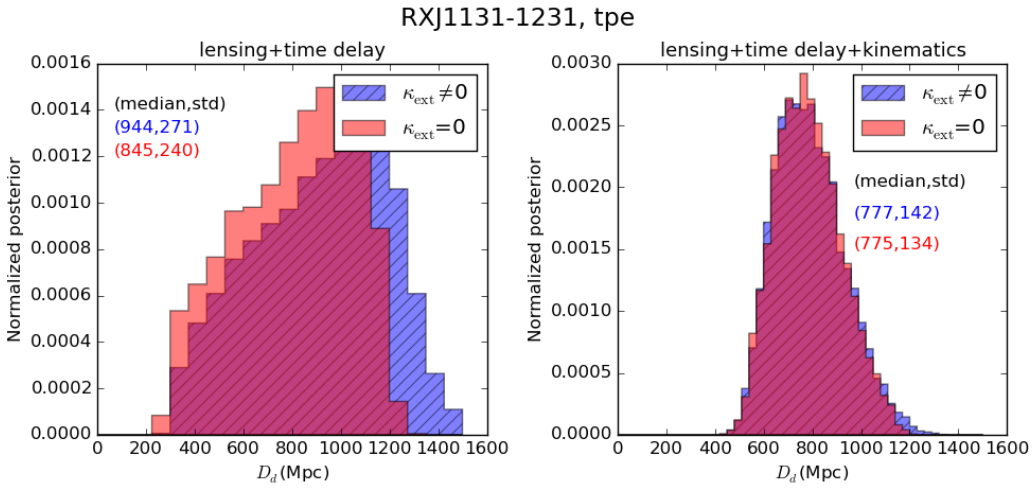

Figure 3.5: The same as figure 3.3, but for TPE anisotropy model. The right panel gives 18% measurement of the angular diameter distance to the lens.

constant. This is the so-called *inverse-distance ladder method* [24]. Using the Joint Light-curve Analysis (JLA, [14]) SNe dataset with two lensing distances we measured as anchors, we constrain the Hubble constant and the SNe nuisance parameters simultaneously. We use Montepython [10] to perform the MCMC analysis. Figure 3.8 shows the resulting cosmological and nuisance parameter constraints from this analysis for flat ΛCDM model. In figure 3.9, we show the Hubble diagram with these parameters. Repeating the same analysis for different cosmological models, we show the constrained Hubble constant using the lensing distances with OM anisotropy, and the JLA for six different cosmological models in 3.10. Depending on the choice of the curvature of the universe, Ω_k, and the equation of state of dark energy, w, the model is either flat ($\Omega_k = 0$) or open ($\Omega_k \neq 0$), and either ΛCDM ($w = -1$), wCDM ($w \neq -1$) or w_aCDM ($w(a) = w_0 + (1+a)w_a$, time-varying equation of state). In case of open curvature ($\Omega_k \neq 0$), the constraints degrades and the Hubble constant is constrained with less precision. However, the constraints on the Hubble constant are nearly independent of the choice of a cosmological model, which shows the robustness of the constraint from two angular diameter distances in combination with the SNe type Ia. We summarize the cosmological and nuisance parameter constrained in tables 3.1 and 3.2, respectively.

Model	Ω_{fld}	Ω_k	h	w_0	w_a
flat ΛCDM	0.702±3.48e-2	≡ 0	0.833±7.90e-2	≡ -1	≡ 0
flat wCDM	0.778±0.122	≡ 0	0.829±8.02e-2	-0.882±0.210	≡ 0
flat w_aCDM	0.782±0.124	≡ 0	0.832±7.91e-2	-0.904±0.214	-1.03±1.39
open ΛCDM	0.559±0.160	0.237±0.258	0.841±8.81e-2	≡ -1	≡ 0
open wCDM	0.419±0.199	0.418±0.287	0.842±8.83e-2	-1.73±0.750	≡ 0
open w_aCDM	0.473±0.222	0.366±0.286	0.841±8.98e-2	-1.46±0.520	5.29e-2±1.20

Table 3.1: Summary of the cosmological parameters for six cosmological models tested.

Model	α	β	M	ΔM
flat ΛCDM	$0.141\pm6.82\times10^{-3}$	$3.11\pm8.30\times10^{-2}$	-18.7±0.206	$-0.0704\pm2.33\times10^{-2}$
flat wCDM	$0.141\pm6.71\times10^{-3}$	$3.11\pm8.30\times10^{-2}$	-18.7±0.210	$-0.0710\pm2.34\times10^{-2}$
flat w_aCDM	$0.141\pm6.79\times10^{-3}$	$3.11\pm8.47\times10^{-2}$	-18.7±0.207	$-0.0710\pm2.30\times10^{-2}$
open ΛCDM	$0.141\pm6.76\times10^{-3}$	$3.11\pm8.38\times10^{-2}$	-18.6±0.229	$-0.0707\pm2.36\times10^{-2}$
open wCDM	$0.141\pm6.75\times10^{-3}$	$3.11\pm8.63\times10^{-2}$	-18.7±0.229	$-0.0705\pm2.37\times10^{-2}$
open w_aCDM	$0.141\pm6.83\times10^{-3}$	$3.11\pm8.31\times10^{-2}$	-18.7±0.232	$-0.0705\pm2.37\times10^{-2}$

Table 3.2: Summary of the nuisance parameters for six cosmological models tested.

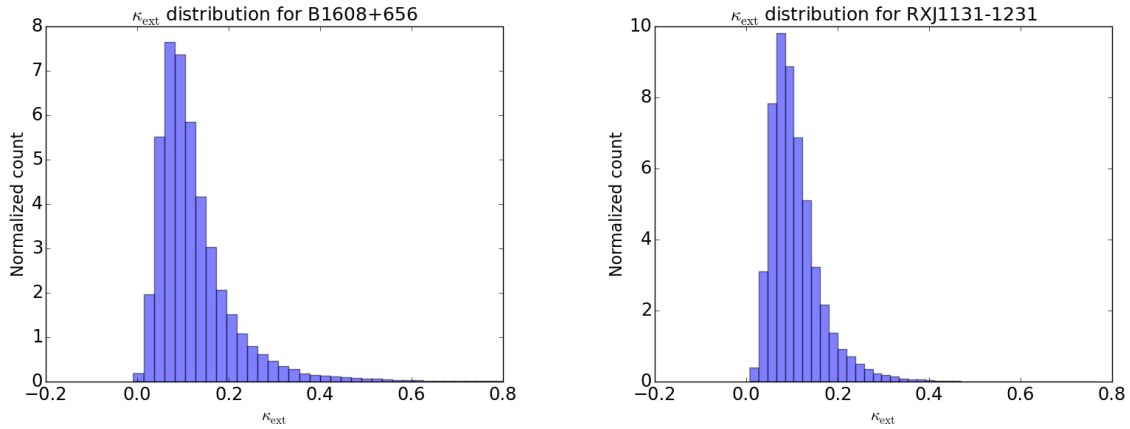

Figure 3.6: The distributions of external convergence $\kappa_{\rm ext}$ used in the analysis for each lens.

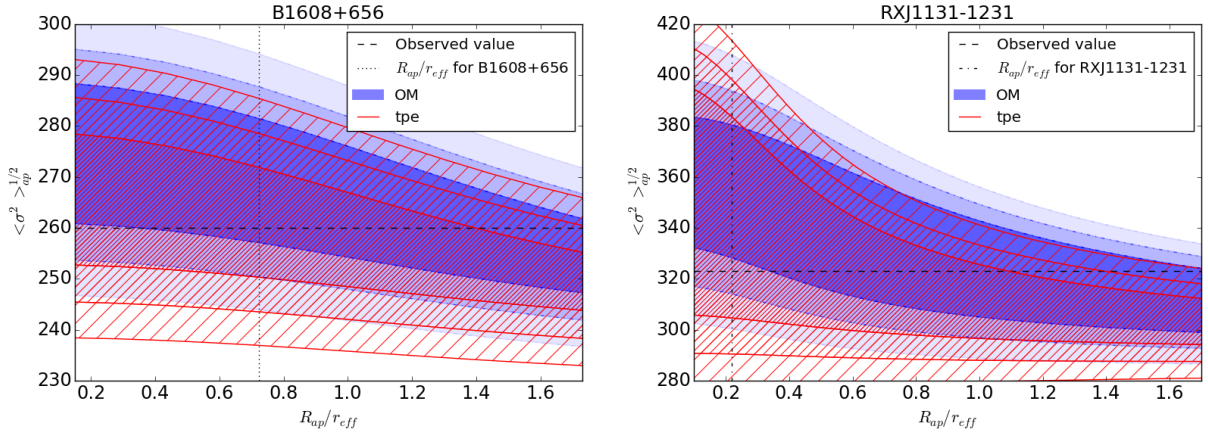

Figure 3.7: The luminosity weighted aperture averaged velocity dispersion for the OM (blue shaded) and TPE (red hatched) anisotropy model, with normalization factors (mass, angular diameter distance, and the Einstein radius) fixed to the posterior value. The left panel is for B1608+656, and the right panel is for RXJ1131-1231. The density profile is constrained tightly by the lensing analysis in range $\gamma = 2.08 \pm 0.03$ for B1608+656 and $\gamma = 1.95 \pm 0.05$ for RXJ1131-1231, respectively, while we give wide range of flat priors on the anisotropy parameters $r_{\rm ani} = [0.5, 5]$ (OM) and $\beta_{\rm in,out}$=[-0.6,0.6] (TPE). From the densest to the least dense shaded areas are for 1, 2, and 3-σ velocity dispersion values, respectively.

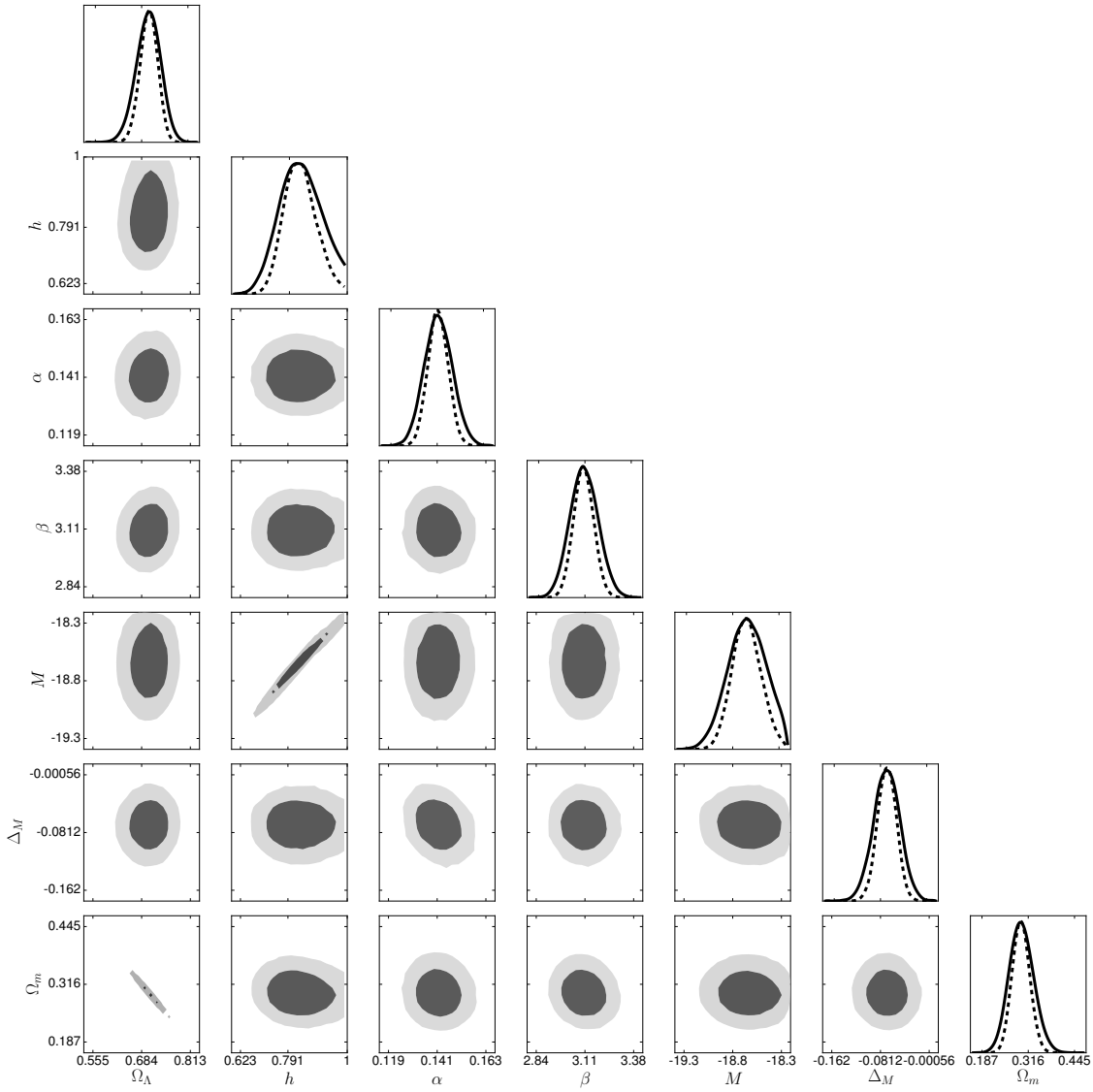

Figure 3.8: One- and two-dimensional posterior distributions of cosmological and nuisance parameters constrained using angular diameter distances from two lenses and JLA, for flat ΛCDM model (solid line). The dark and light shaded area in two-dimensional contours are for 68% and 95% Confidence Level. The dashed lines are for the average likelihood distribution.

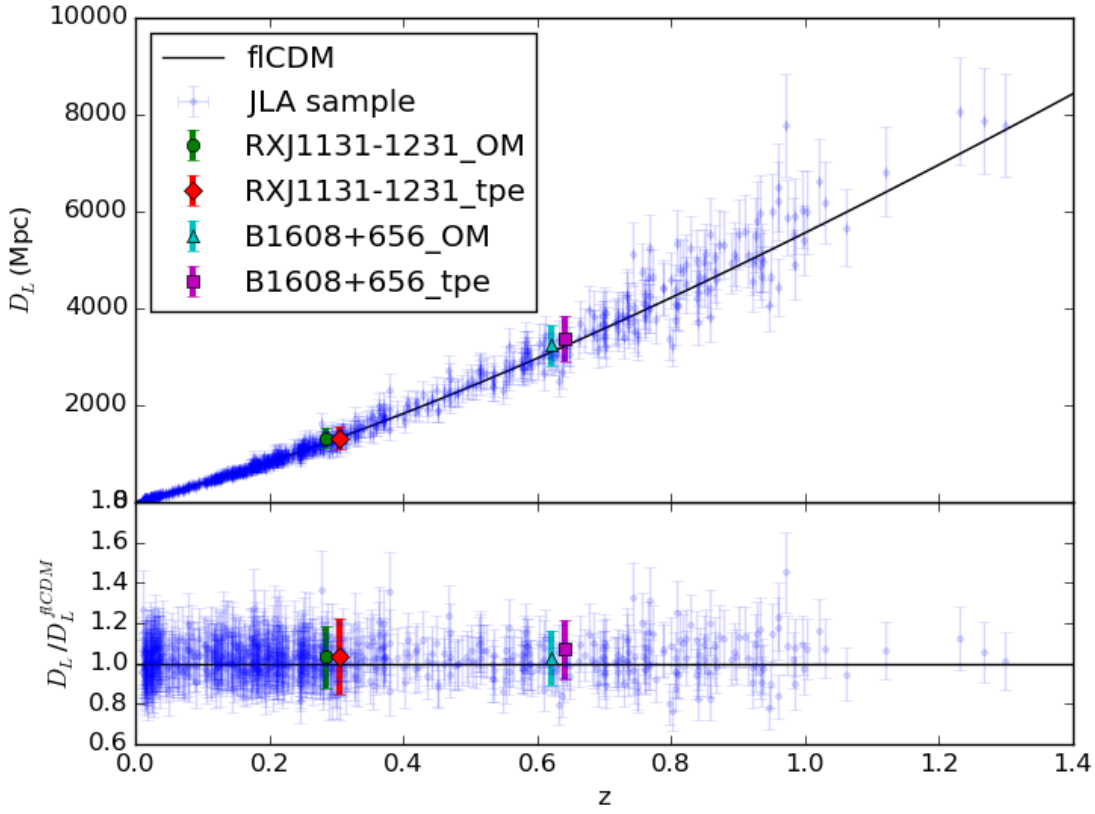

Figure 3.9: The Hubble diagram with 740 SNe from JLA normalized by two lensing distances. We used the distances constrained from OM anisotropy model (green circle and cyan triangle points). We overplot the distances constrained from TPE model (red diamond and purple square points) for comparison. The solid line is flat ΛCDM model fitted with the cosmological and nuisance parameters constrained from figure 3.8.

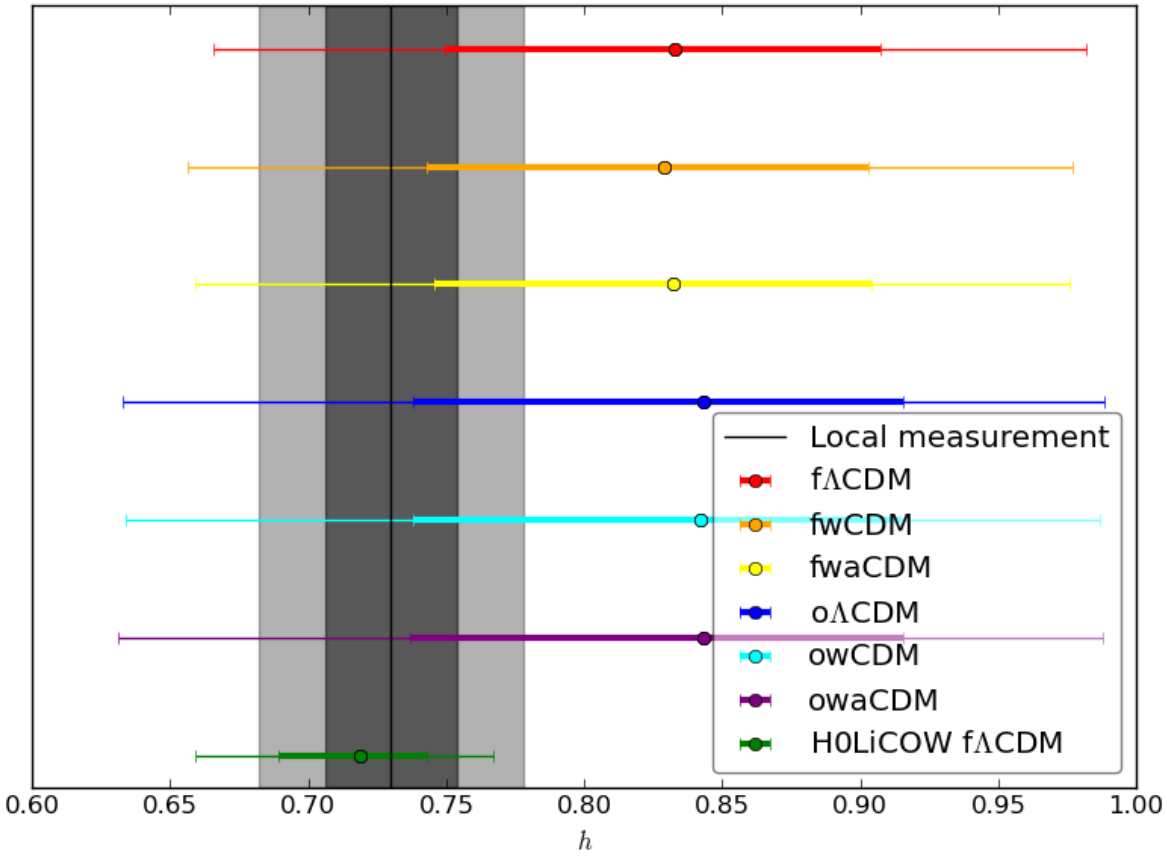

Figure 3.10: The Hubble constant constrained from six cosmological models we tested. The gray shaded area is the constraints using local distance measurements [87], and the green line is from three time-delay distances measured by H0LiCOW collaboration [16]. The thick and thin solid lines are where the 68% and 95% Confidence Levels are.

Chapter 4

Cosmological Implications from Future Observations

We use predictions for detecting quadruply lensed quasars to calculate the expected cosmological constraints from lensing distances in future large-scale surveys. We compare the constraints with other distance probes, in particular, type Ia supernovae and Baryon Acoustic Oscillations (BAO). The contents are published in [51].

4.1 Method

Each well-modeled time-delay lens system yields two distance(-like) quantities, $D_A(EL)$ and $D_{\Delta t}$. The uncertainties of $D_A(EL)$ and $D_{\Delta t}$ are dominated by the velocity dispersion and the external convergence, respectively. In this work, we make an optimistic assumption that we can measure both distances with 5% uncertainties, which requires a few per cent measurement of the spatially resolved velocity dispersion of the lens galaxy, as well as a good understanding of the mass distribution along the line-of-sight, that is obtainable by simulations and observations of the lens environment [100, 50]. Regarding the lens mass model, the power-law density profile in ref. [50] is widely used due to its ability to fit the imaging data near the image positions. The local density profile is well reconstructed with the model if the images are spatially extended such that information from thousands of intensity pixels can be used. However, ref. [91, 92] have pointed out that the information obtained by the lensed images cannot uniquely determine the shape of lens mass profile due to the so-called *Source-Position Transformation* (SPT). Specifically, they focused on the degeneracy between composite density profiles and a power-law mass profile, and have shown that fixing the shape of the lens mass profile as a power law can break the SPT. However, they have also mentioned that these models can be distinguished if more information is available: for example, if more than three images with time delays are observed, the degeneracy can be broken as the general SPT does not conserve the time delay ratios. In ref. [100], the robustness of the measured time-delay distance is tested with power law and composite model under the presence of lens kinematics data and shown to be nearly

independent of the choice of the model. Ref. [31] has shown that the so-called *Mass-Sheet Transformation* (MST), which is a special case of the SPT, scales the time delays by the same factor, and thus conserves the time-delay ratio, can bias the mass modeling. However, ref. [114] has empirically shown using the Illustris simulation that most of the early-type galaxies with high ($\sigma > 200\text{kms}^{-1}$) velocity dispersion, which most of the lens galaxies are, show nearly power-law behavior. The MST degeneracy can also be broken if additional information on the lens galaxy (e.g. velocity dispersion) is obtained. Ref. [52] has shown that the existence of substructures in the lens galaxy perturbs the time delay: However, the effect of perturbation on the time delays (<1 day) is typically smaller than the currently available time-delay measurement uncertainties, and thus both distances are mainly determined by the global mass distribution rather than the substructures. Thus we claim that these precision measurements on both distances are possible, but only when good quality imaging / kinematics data as well as time-delay measurements are available. The correlation between the two distances is negligible, because the uncertainties in the velocity dispersion and the external convergence are uncorrelated. For completeness, and to study the case in which 5% measurements are not achievable, we also quote the constraints assuming the 10% measurements in the parentheses.

To study the expected cosmological constraints from lenses, we need to specify the distribution of lens and source redshifts. We use the catalog of mock lenses in ref. [78] to obtain the redshift distribution of time-delay lenses with double and quadruple images expected for the Large Synoptic Survey Telescope (LSST) [2, 48]. Although LSST itself is expected to find around ten thousand lensed quasars, there are only about 400 systems that would have good time delay measurements [62]. To obtain distances from a lens system with a reasonable accuracy, a good mass model of the lens galaxy is also required, as both the time-delay distance and the angular diameter distance are sensitive to the mass distribution of the lens. Ancillary data both in terms of high-resolution imaging and spectroscopy of the lens systems are needed for accurate lens mass modeling. Therefore, we select lens systems from the mock catalog with the following criteria for acquiring ancillary data with relative ease: (1) the quasar image separation is $> 1''$, (2) the third brightest quasar image has an i-band magnitude $m_i < 21$, and (3) the lens galaxy has $m_i < 22$. The criteria on the quasar image separation and brightness make it easier to measure the time delays with high precision (uncertainty of a few percent). Furthermore, a sufficiently wide quasar image separation is required for extracting the Einstein ring of the quasar host galaxy and measuring the lens velocity dispersion for mass modeling. The lens galaxy also needs to be of sufficient brightness for measuring the lens velocity dispersion. In this work, we focus on quadruply imaged lens systems as they provide more information than doubly imaged systems. After applying these criteria to the LSST mock lens sample, we obtain 55 quadruple lens systems as the best cases of obtaining ancillary data.

Figure 4.1 shows the source and the lens redshift distribution of quadruple lenses in our mock catalog. For the visualization purpose, only in this figure, the total number of detectable lenses is oversampled by a factor of 5 (based on the catalog from ref. [78]) to populate the histogram. The expected constraints reported in this paper are derived from the actual distribution of the 55 lenses.

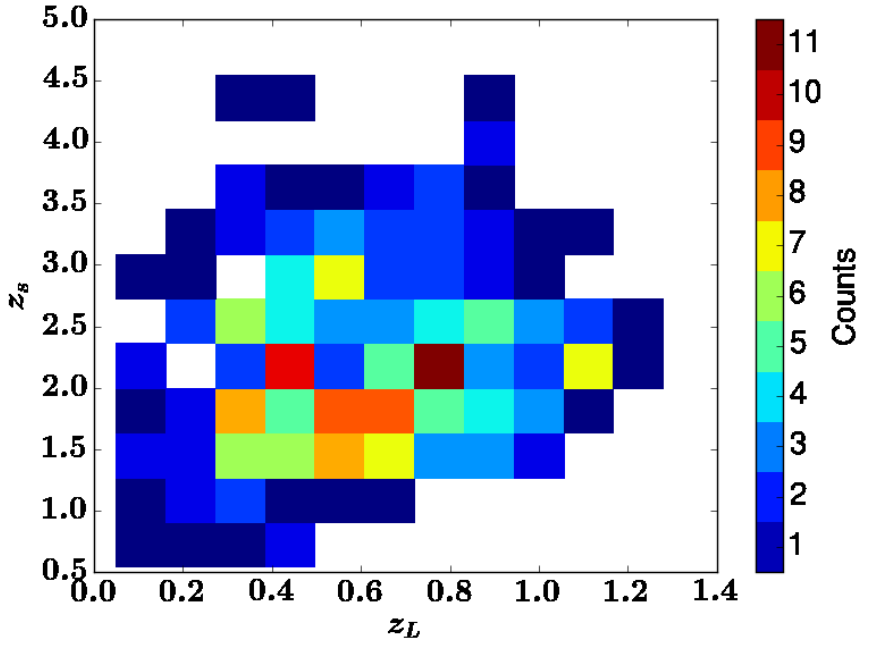

Figure 4.1: The distribution of source (z_S) and lens (z_L) redshifts of quadruply imaged time-delay lenses expected for LSST [78]. These lens systems are the best ones to obtain ancillary data and measure distances with required precision. For visualization purposes, the number of the lenses is oversampled by a factor of five only in this figure (thus each pixel does not necessarily have a multiple-of-5 value).

Since these are the bright lens systems, a fraction of these systems will already be discovered in the current imaging surveys. In particular, we expect that ∼25% of these systems will be discovered in the Dark Energy Survey (DES)[1] and the Hyper Suprime-Cam (HSC; [71])[2] Survey. Furthermore, we expect a few more quadruple lens systems from the northern areas of the HSC Survey that are not covered in DES and LSST. Therefore, even though we focus here on the LSST sample, our cosmographic predictions are also relevant for the upcoming years before the LSST era as new lens systems in the current imaging surveys are discovered and monitored.

We explore constraints on two variations of ΛCDM model. Both assume a curved universe ($\Omega_k \neq 0$) and an unknown equation of state of dark energy ($w \neq -1$). The first model assumes that w is a constant (owCDM) with the following cosmological parameters:

$$\vec{\theta} \in \{\Omega_m, \Omega_k, w, h\} \quad (\text{o}w\text{CDM model}). \tag{4.1}$$

The second model further assumes that w varies in time with $w = w_0 + (1 - a)w_a$ [63]

[1] http://www.darkenergysurvey.org/index.shtml
[2] http://www.naoj.org/Projects/HSC/surveyplan.html

(ow_zCDM):

$$\vec{\theta} \in \{\Omega_m, \Omega_k, w_0, w_a, h\} \quad (ow_z\text{CDM model}). \tag{4.2}$$

We choose the fiducial cosmology following Planck 2015 ($\Omega_m = 0.308$, $\Omega_k = 0$, $w_0 = -1$, $w_a = 0$, and $h = 0.678$) [3].

We use the Fisher information matrix (hereafter the Fisher matrix) to calculate the constraining power of the cosmological probes. For a data vector $\vec{d}(\vec{\theta})$ with a set of parameters $\vec{\theta}$, the Fisher matrix $\mathbf{F}$ is given by

$$F_{ij} = \sum_{\alpha\beta} \frac{\partial d_\alpha}{\partial \theta_i} \, \mathbf{Cov}^{-1}_{\alpha\beta} \, \frac{\partial d_\beta}{\partial \theta_j} \ , \tag{4.3}$$

where indices α and β run over the observables, and $\mathbf{Cov}(\vec{d})$ is the data covariance matrix. For lenses, $\mathbf{Cov}_{\alpha\beta} = \delta_{\alpha\beta}\sigma_\alpha^{-2}$, where $\delta_{\alpha\beta}$ is the Kronecker delta as we assume no correlation between the different lens systems and between the two measured lensing distances D_A and $D_{\Delta t}$ of each lens system. The uncertainty in each distance σ_α is $\sigma_\alpha = 0.05 d_\alpha$ ($\sigma_\alpha = 0.1 d_\alpha$) as we assume 5% (10%) precision measurements of both distances. The inverse of the Fisher matrix, $\mathbf{F}^{-1}$, gives the covariance matrix of the parameters, and the marginalized uncertainty on the i-th parameter is calculated as $(\mathbf{F}^{-1})^{1/2}_{ii}$. If the normalization of $\mathbf{F}$ increases by a factor of n, then the normalization of the parameter covariance matrix decreases by a factor of n, thus the error bar on each parameter tightens by a factor of $\sqrt{n}$.

4.2 Single-probe constraints combined with the Planck distance prior

In this section we investigate the constraining power of the time-delay lenses expected from LSST in section 4.2.1, and compare the constraints to that of two other cosmological distance probes, the BAO data from Baryon Oscillation Spectroscopic Survey (BOSS) data release (DR) 11 in section 4.2.2, and the SNe data from Joint Light-curve Analysis (JLA) in section 4.2.3. We examine the future prospects of constraints from BAO and SNe using LSST data in section 4.2.4. We combine each probe with the CMB distance prior calculated from the Planck 2015 result [4]. The distance prior is calculated using the shift parameter, $R_{\rm shift}$, and the multipole corresponding to the sound horizon at the moment of last scattering, l_*. The definitions of these parameters are

$$\begin{aligned} R_{\rm shift} &\equiv \sqrt{\Omega_m H_0^2} D_A(z_*)/c, \\ l_* &\equiv \pi \frac{D_A(z_*)}{r_s(z_*)}, \end{aligned} \tag{4.4}$$

where $z_* = 1089.94$ is the redshift of the last scattering surface, and $r_s(z_*) = 144.89$ Mpc is the size of the sound horizon at redshift z_*. The distance prior compresses information in the CMB power spectrum relevant for dark energy to two numbers.

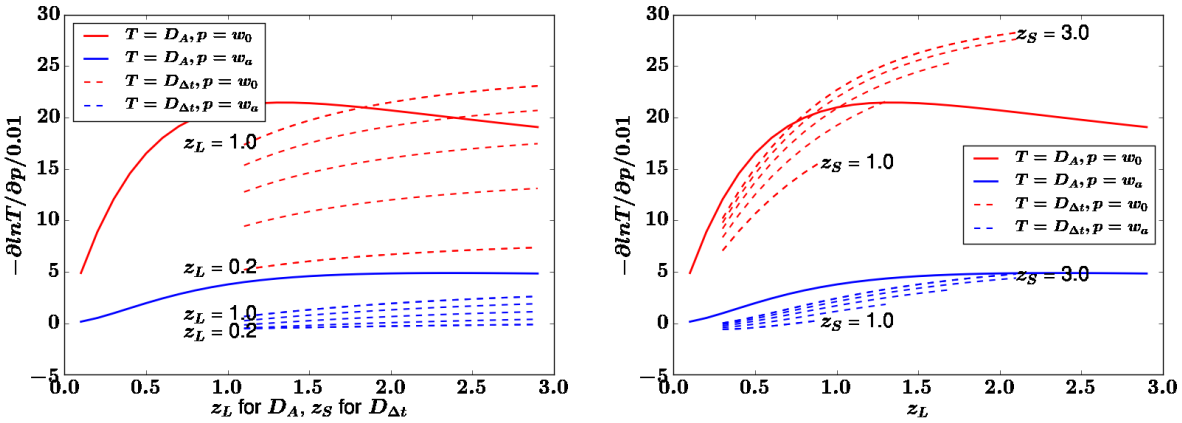

Figure 4.2: Logarithmic derivatives of distances $T = (D_A, D_{\Delta t})$ with respect to the cosmological parameters $p = (w_0, w_a)$, as a function of redshift. Left panel: The solid lines show $T = D_A$, while the dashed lines show $T = D_{\Delta t}$, with various combinations of the lens and the source redshift. Each dashed line corresponds to one z_L in range [0.2,1.0] in increments of 0.2, and shows $-\partial \ln D_{\Delta t}/\partial p/0.01$ as a function of z_S. We only show z_S which is higher than the highest z_L in this range. Right panel: The solid lines show $-\partial \ln D_A/\partial p/0.01$ as a function of z_L. Each dashed line corresponds to one z_S in range [1.0,3.0] in increments of 0.5, and shows $-\partial \ln D_{\Delta t}/\partial p/0.01$ as a function of z_L. We only show z_L which is lower than the lowest z_S in this range. Both panels show that D_A is always more informative than $D_{\Delta t}$ for constraining w_a (i.e., $|\partial \ln D_A/\partial w_a| > |\partial \ln D_{\Delta t}/\partial w_a|$), and D_A is often more informative than $D_{\Delta t}$ on w_0.

4.2.1 Time-delay lenses

We first show how sensitive $D_{\Delta t}$ and $D_{\rm A}(EL)$ are to w_0 and $w_{\rm a}$ as a function of z_L and z_S.

In figure 4.2, we show $\partial \ln T/\partial p$ (where $T = (D_A, D_{\Delta t})$ and $p = (w_0, w_a)$). The larger the absolute value of $\partial \ln T/\partial p$ is, the bigger the unmarginalized sensitivity of the distance T becomes to the parameter p, with all the other cosmological parameters fixed at the fiducial values. Also, equation 4.3 shows that the information is proportional to $\partial T/\partial p$. The higher z_L is, the more sensitive $D_{\Delta t}$ becomes to both w_0 and w_a for a given z_S, and vice versa. We find that D_A is always more sensitive to w_a, when all the other parameters are fixed at the fiducial values.

In figure 4.3, we show the expected 1-σ uncertainties in w for the owCDM model and w_0 for the ow_zCDM model (combined with the Planck distance priors), with all the other parameters marginalized over. As the Fisher matrix is proportional to $n_{\rm quads}$, the marginalized uncertainty in $D_A(EL) + D_{\Delta t}$ scales as $\propto 1/\sqrt{n_{\rm quads}}$. For o$w$CDM (left panel of figure 4.3), the Planck distance priors combined with either D_A (blue dashed line) or $D_{\Delta t}$ (red dotted) from lenses improve the constraint on w significantly compared to the lens-only case (black dash-dot). Combining all improves the constraint further by 30%,

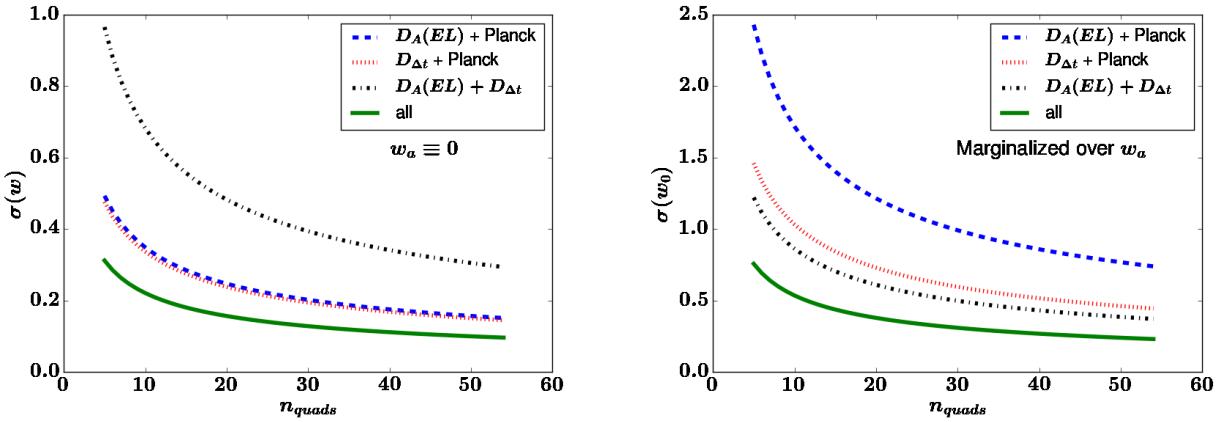

Figure 4.3: The 1-σ uncertainty in w, denoted as $\sigma(w)$, and that in w_0, denoted as $\sigma(w_0)$, from time-delay lenses as a function of the number of lenses for the left and the right panel, respectively. The black dash-dot line is the lens-only data, while the other lines use the Planck distance priors combined with D_A (blue dashed), $D_{\Delta t}$ (red dotted), or both (green) from lenses. (Left) owCDM model. (Right) ow_zCDM model marginalized over w_a, as well as all the other parameters.

in comparison to that of the $D_{\Delta t}$ + Planck. For ow_zCDM (right panel of figure 4.3), we find that the lens-only (D_A + $D_{\Delta t}$) breaks the degeneracy between w_0 and the other parameters more efficiently than either combination of Planck + D_A or Planck + $D_{\Delta t}$, yielding a tighter constraint. Combining all improves the constraint further by a factor of 2 in comparison to $D_{\Delta t}$ + Planck. In case of 10% uncertainties on both distances, the fractional improvement due to the addition of D_A to Planck + $D_{\Delta t}$ is the same as that for 5% uncertainties, but the overall constraints become weaker by a factor of 2. In other words, the marginalized uncertainties in w_0 and w_a are proportional to the uncertainties in the lensing distance measurements, as the Planck data alone are unable to constrain w_0 and w_a.

The left panel of figure 4.4 is useful for understanding these results. The marginalized uncertainties in w_a from either D_A or $D_{\Delta t}$ individually are similar. The Planck distance prior (the black line) provides a degenerate combination of w_0 and w_a, thus cannot be marginalized. However, it is nicely orthogonal to the ones from D_A (the blue contour) and $D_{\Delta t}$ (the red contour). Thus, including the combination of D_A and Planck with the previous lensing constraints from $D_{\Delta t}$ reduces the uncertainty in w_a significantly. We also note that the constraint on w_a from D_A + $D_{\Delta t}$ (the green contour) is significantly tighter than the naive addition of the blue and the red contour, which indicates that this combination of distances effectively breaks the degeneracy between the equation of state of dark energy and the other parameters over which we marginalize.

Next, we study the degeneracy structure of parameters constrained from lensing distances, by shifting a parameter and calculating the response of the other parameters to the shift in order for the likelihood to be maximized. Specifically, we describe the degeneracy

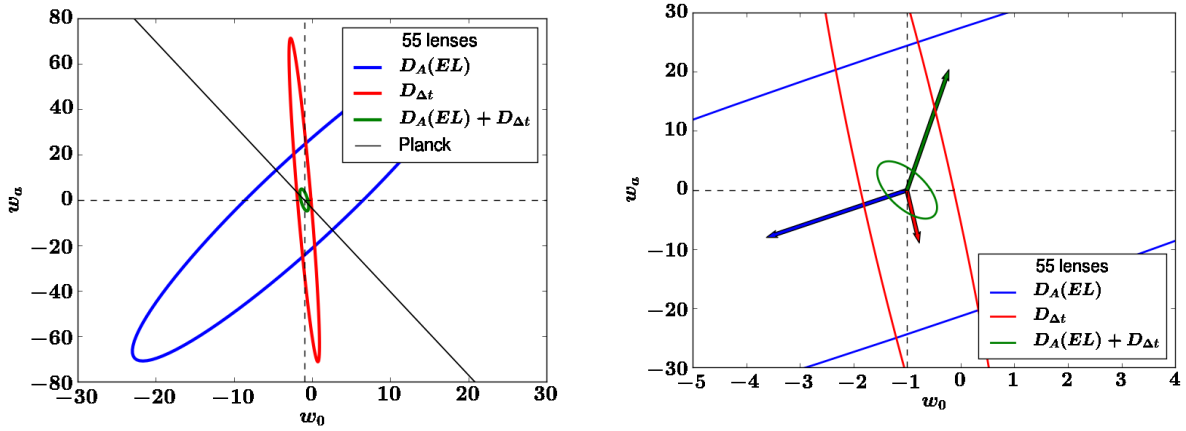

Figure 4.4: (Left) The 68 per cent CL in the w_0-w_a plane constrained from 55 lenses for the ow_zCDM model. The blue line is the marginalized constraint from D_A, the red line is the marginalized constraint from $D_{\Delta t}$, and the green line is the marginalized constraint from the combination of the D_A and $D_{\Delta t}$. The black line is the unmarginalized constraint from Planck only. The horizontal and vertical dashed lines correspond to the fiducial values of $w_0 = -1$ and $w_a = 0$, respectively. (Right) Zoom in of the left panel, with degeneracy directions plotted as arrows. The arrows show the response of w_0 and w_a to the shift $\Delta\Omega_k = 0.02$. The blue arrow is the response for $D_A(EL)$, the red arrow is for $D_{\Delta t}$, and the green arrow is for the combination of two. The larger the arrow is relative to the contour, the more sensitive the distance is to Ω_k. The alignment between the arrows and the major axis of the contours indicate that the constraints from D_A (the red arrow and the red contour) and $D_{\Delta t}$ (the blue arrow and the blue contour) are individually dominated by Ω_k. However, the degeneracy between Ω_k and w is shown to be broken when D_A and $D_{\Delta t}$ are combined. For visualization purposes, the sizes of the arrows are inflated by a factor of 100.

between the curvature density Ω_k and the equation of state parameters w_0 and w_a. When Ω_k is shifted by an amount $\Delta\Omega_k$, to maximize the likelihood at the new fiducial value $\Omega_k + \Delta\Omega_k$, all the other parameters have to be shifted accordingly. The general expression for the shift in an arbitrary parameter $\Delta\theta_i$ due to a shift in a fixed, single parameter $\Delta\theta_k$ that maximizes the likelihood can be calculated as

$$\Delta\theta_i = \sum_k \Delta\theta_k \frac{(\mathbf{F}^{-1})_{ik}}{(\mathbf{F}^{-1})_{kk}}, \tag{4.5}$$

where $\mathbf{F}^{-1}$ is the inverse of the Fisher matrix. In figure 4.4, the right panel shows the projection of the shift vectors to w_0-w_a plane as arrows, along with the marginalized constraint contour at the fiducial parameter to display it quantitatively. For each of D_A and $D_{\Delta t}$, the degeneracy directions are parallel to the major axes of the contours, which indicates that the degeneracies with Ω_k dominate the dependences of w_0 and w_a to other parameters. However, when the two distances are combined, the curvature degeneracy

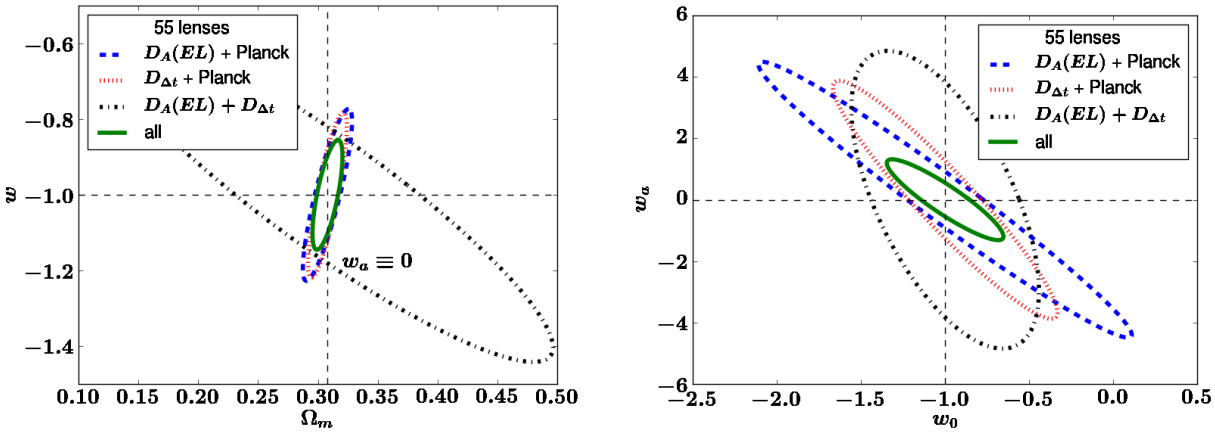

Figure 4.5: The marginalized 68 per cent CL constraints from 55 lenses in the (left) Ω_m-w plane for the owCDM model, and (right) w_0-w_a plane for the ow_zCDM model. The black dash-dot lines show the lens-only constraints from D_A + $D_{\Delta t}$, the blue dashed lines the constraints from D_A + Planck, the red dotted lines the constraints from $D_{\Delta t}$ +Planck, and the green solid lines the combination of the two distances + Planck.

breaks and the alignment between the error contour and the shift disappears (the green arrow and the green contour are not aligned). The relative size of the arrow to the contour shows the sensitivity of the probe to Ω_k: the bigger the vector is with respect to the contour, the more sensitive the probe is to the change in Ω_k. By comparing the relative size of the red and the blue arrows to the red and the blue contours, we show that D_A and $D_{\Delta t}$ are comparably sensitive to $\Delta\Omega_k$, but the combination of two increases the sensitivity significantly (the green arrow and the green contour).

The left panel of figure 4.5 shows the joint constraints on Ω_m and w for owCDM. We find that the Planck distance prior plays an important role in constraining Ω_m, while using both D_A and $D_{\Delta t}$ combined with Planck distance prior improves the constraint on w by about 30% compared to the case of $D_{\Delta t}$ combined with Planck distance prior. The right panel of figure 4.5 shows the same for w_0 and w_a for ow_zCDM. We also show the lensing distances + Planck constraint on H_0 in Appendix B.

Next, we compare these constraints with those from Planck + BAO and Planck + SNe. We calculate the constraints from BAO and SNe using currently available data (BOSS DR11 for BAO, JLA sample for SNe).

4.2.2 BAO

BOSS DR11 provides the volume-averaged distance, $D_V \equiv (cz(1+z)^2 D_A^2/H)^{1/3}$, at two effective redshifts (0.32, 0.57) obtained from the BAO peak position in the spherically averaged two-point functions. The lower redshift is the LOWZ sample, and the higher redshift is the CMASS sample. Also, by separately measuring the two point functions along the line of sight and the direction perpendicular to it, the DR11 CMASS sample

separately constrains the angular diameter distance D_A and the Hubble parameter H at $z = 0.57$ [9, 90]. To account for the correlation among D_V, D_A and H, we use the full likelihood of the CMASS sample for the analysis provided by the BOSS collaboration [9].

In our analysis we assume that the sound horizon scale at the baryon drag epoch, $r_{s,drag}$, is fixed as $r_{s,drag} = 149.28$ Mpc [9]. We then combine the cosmological constraints from D_V at $z = 0.32$, and D_A and H at $z = 0.57$ with the Planck distance prior. We calculate the Fisher matrix by taking the derivatives of the log likelihood at the fiducial cosmology. The results are shown in figure 4.6. The precision of the BAO data yields the narrowest contours on the Ω_m-w (for owCDM) and w_0-w_a (for ow_zCDM) planes. However, due to the limited number of redshifts ($z = 0.35$ and $z = 0.57$), the degeneracy is not broken efficiently; thus, the expected Planck + lens from 55 lenses can improve the constraints significantly, even though the precision of lensing data per redshift is not as precise as BAO.

4.2.3 SNe

We now study the constraints from Planck + SNe. We use the JLA data [14] to calculate the constraints from SNe. JLA uses Supernovae Legacy Survey (SNLS), Sloan Digital Sky Suvey-II (SDSS-II) Supernova survey and a few low-redshift samples. The redshift of subsamples are: the low-redshift sample ($z < 0.1$), SDSS-II ($0.05 < z < 0.4$), and SNLS ($0.2 < z < 1$). There are 740 spectroscopically confirmed type Ia SNe in JLA. SDSS-II is used for anchoring the distances, and also an empirical relation between the host galaxies and the supernovae brightness is used as an extra calibration for the absolute magnitude of the SNe. For the calibration, there are 4 additional nuisance parameters that are taken into account in JLA: α, which scales the stretch of the light curve in time-domain; β, which scales the color at the peak of the light curve; M, which is the absolute B band magnitude of the SNe at the peak of the light curve; and Δ_M, which characterizes the peak absolute magnitude change due to the stellar mass of the host galaxy.

We use Montepython [10] to sample the JLA likelihood. Specifically, we run Markov Chain Monte Carlo to sample the likelihood surface, and compute the covariance matrix in the cosmological parameters. We then use its inverse as the JLA Fisher matrix. The results are shown in figure 4.6. While the absolute distances, such as those from BAO and lenses, are effective at measuring Ω_k when combined with CMB [53], the relative distances from SNe are not. Thus, when Ω_k is set free, the constraints on Ω_m for owCDM from 55 lenses combined with Planck are significantly better than those from 740 SNe combined with Planck.

4.2.4 Comparison to future BAO and SNe predictions

With several billion galaxies expected to be detected with LSST, BAO will allow measurements of distances with $\sim$2% precision in the redshift range $1 < z < 3$ [115]. Combined with Planck, BAO will constrain w_0 with uncertainty $\sim$0.4 and w_a with $\sim$1 [1]. Also, 500,000 SNe are expected to be detected in 10 years of LSST operation in the redshift

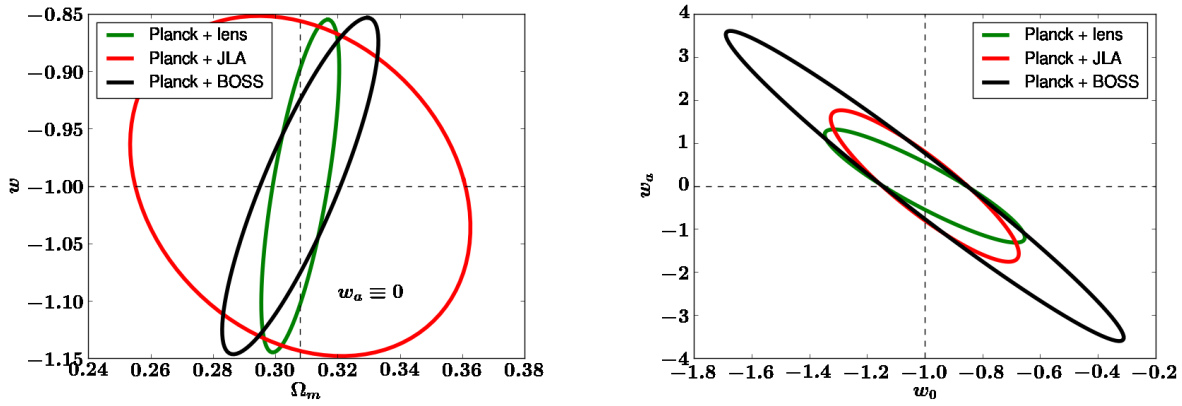

Figure 4.6: The marginalized 68 per cent CL constraints from strong lenses, SNe, and BAO, each combined with the Planck distance prior, in the (left) Ω_m-w plane for the owCDM model, and (right) w_0-w_a for the ow_zCDM model. The green lines show the constraints from Planck + lens, the red lines Planck + JLA, and the black lines Planck + BOSS.

range $0.1 < z < 1.2$. With a subsample of 50,000 SNe only, the data will constrain w_0 with uncertainty $\sim$0.05, and w_a to order unity, assuming a flat universe [2]; in combination with Planck, the full sample of SNe constraints will be $\sim$0.25 for w_0 and $\sim$1.2 for w_a for the ow_zCDM model [1]. We note that a modest sample of 55 lenses combined with the Planck distance prior constrains w_0 and w_a to $\sim$0.4 ($\sim$0.7 for the case of 10% uncertainties on the lensing distances) and $\sim$1.2 ($\sim$2.5), respectively (see, e.g., Figure 6), which is comparable in precision to those expected from future BAO or SNe samples in the LSST era, when each is combined with Planck. Therefore, strong lenses provide an independent and competitive probe of dark energy. Needless to say, lensing, SNe, and BAO are affected by different systematic errors, and thus cross-checking the results using these three low-redshift probes of the expansion of the universe will be powerful.

4.3 Pivot redshift

The equation of state of dark energy, $w(z)$, can be re-written as

$$w = w_0 + (1-a)w_a = w_p + (a_p - a)w_a, \tag{4.6}$$

where $w_p \equiv w_0 + (1-a_p)w_a$ [44]. In this parameterization, the *pivot redshift* $z_p = \frac{1}{a_p} - 1$ is defined as the redshift where the uncertainty in w is minimized. The uncertainty in w_p shows how well a probe can measure the equation of state w, as w_p is orthogonal to w_a by construction, and thus is not coupled to the time variation of w [44, 8]. z_p shows at which redshift the main constraint on w is coming from: this pivot redshift varies depending on the probes, their redshift distributions and the measurement uncertainties, and can be negative. Figure 4.7 shows the constraints on (w_p,w_a) using the lens distances alone in

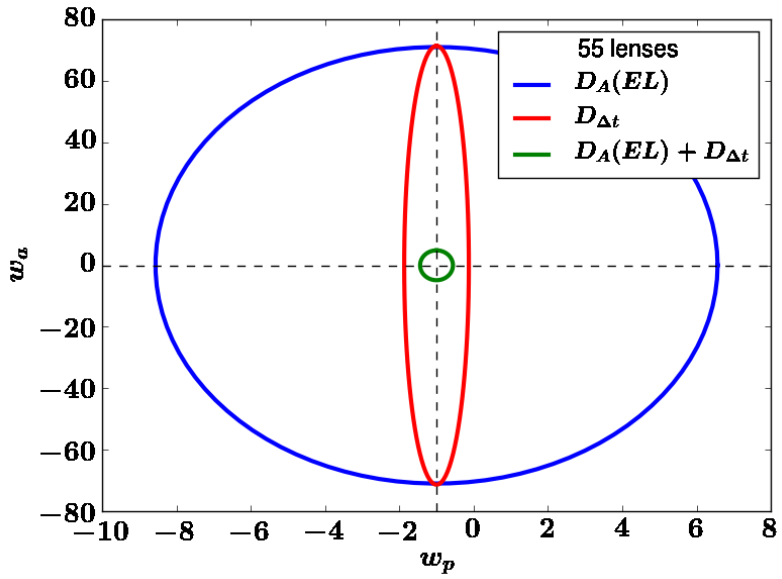

Figure 4.7: Same as figure 4.4, but in w_p-w_a plane. z_p is calculated separately for each probe: $z_p(D_A) = -0.226$, $z_p(D_{\Delta t}) = 0.0251$, and $z_p(D_A + D_{\Delta t}) = 0.0759$.

w_p-w_a plane. The constraint on w_p from D_A alone is much weaker than that from $D_{\Delta t}$, while the constraints on w_a from both distances are comparable. However, by combining the two distance measures, the constraint on w_p improves by a factor of 2, and that on w_a improves by a factor of 24, due to complementary degeneracy directions as shown in figure 4.4.

The pivot redshift, z_p, and the uncertainties in w_p and w_a for two combinations of probes (Planck + BOSS + JLA, and 55 lenses (denoted by Lens) + Planck + BOSS + JLA) for three different cosmological cases ((1) $\Omega_k = 0$ and $\Omega_m = 0.309$, (2) $\Omega_k = 0$ and marginalized over Ω_m, and (3) marginalized over both Ω_k and Ω_m) are summarized in table 4.1. In comparison to Planck + BOSS + JLA, adding 55 lenses tightens the FoM by a factor of 2.78 (1.62), 1.79 (1.45) and 2.17 (1.41) for the three models, respectively, given 5% (10%) distance measurements. In particular, when Ω_k is allowed to vary, the uncertainty in w_a is reduced by almost a factor of 2 by including the lenses, which shows that the combination of D_A and $D_{\Delta t}$ is powerful in breaking the Ω_k-w degeneracy. This is consistent with our argument in section 4.2.1. Also z_p becomes higher as we include the lens distances in every case, which is typically beneficial when combining these probes of geometry with probes of the growth of cosmic structure.

Cosmological model	Probe	z_p	$\sigma(w_p)$	$\sigma(w_a)$	FoM
Flat, fixed Ω_m	CMB+BAO+SNe	0.288	0.0375	0.364	73.3
	Lens+CMB+BAO+SNe	0.487	0.0296	0.166	204
Flat, marginalized over Ω_m	CMB+BAO+SNe	0.358	0.0487	0.413	49.7
	Lens+CMB+BAO+SNe	0.386	0.0310	0.363	88.9
Marginalized over Ω_m and Ω_k (ow_zCDM)	CMB+BAO+SNe	0.215	0.0625	1.03	15.5
	Lens+CMB+BAO+SNe	0.245	0.0479	0.621	33.6

Table 4.1: Pivot redshift, z_p, uncertainties in w_p and w_a, and the Figure of Merit (FoM) for Planck + BOSS + JLA and Lens + Planck + BOSS + JLA. We test three cases: (1) $\Omega_k = 0$ and $\Omega_m = 0.309$, (2) $\Omega_k = 0$ and marginalized over Ω_m, and (3) marginalized over both Ω_k and Ω_m. As there is no correlation between w_p and w_a, FoM $= 1/(\sigma(w_p)\sigma(w_a))$.

4.4 Conclusion

There is more valuable cosmological information in the strongly lensed systems than measurements of the Hubble constant from time delays. In this chapter, we have demonstrated that the addition of the angular diameter distance measurements to the quantity ($D_{\Delta t}$) that captures the cosmological information from time delays in the same sample of lenses provides crucial help in breaking cosmological parameter degeneracies. This improvement is most significant in some of the most interesting parameterizations that are currently being studied, such as when curvature of the universe and the time-variation in the equation of state of dark energy are allowed to be nonzero.

We have calculated the lensing constraints based on the predictions for the LSST survey, adopting a catalog of 55 quadruply imaged lenses (out of a much larger total number) that should have sufficiently good information that all observable quantities of interest in the lenses can be accurately measured. We have combined the forecasted lensing information from both the angular diameter distance and the time-delay distance. We then compared this lensing constraints with that from the BOSS DR11 and from the JLA type Ia supernova sample – each combined with the Planck 2015 distance prior.

We find that the combined lensing information significantly helps constrain the cosmological parameters, particularly when curvature is allowed to vary and when the equation of state of dark energy is allowed to be time-dependent. For example, 5% (10%) precision measurement of lensing distances would improve the current BAO+CMB+SN constraints on w_a by a factor of two (20%), and those on the overall figure of merit of dark energy by about a factor of two (50%) relative to the case with no lensing (see Table 4.1 and figure 4.8). Key to this significant improvement is lensing's ability to break the degeneracy between curvature and the equation of state parameters; see figure 4.4.

We are therefore very optimistic about the prospects of a select, accurately observed subsample of strong gravitational lenses to improve our constraints on dark energy. Fortunately, the lensing samples are a guaranteed product of the current and upcoming wide-

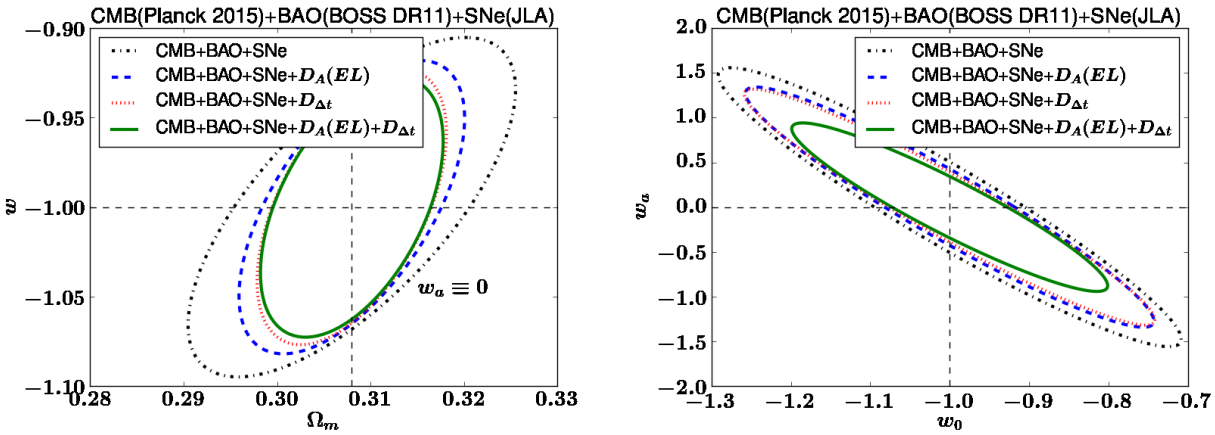

Figure 4.8: The marginalized 68 per cent CL from strong lenses, combined with Planck, BAO and SNe in the (left) Ω_m-w plane for the owCDM model, (right) w_0-w_a for the ow_zCDM model.

field, deep surveys such as HSC, DES and LSST.

Chapter 5

Current & Future Perspectives: Challenges in Strong Lensing Cosmography

I have shown that strong lensing is a prominent tool in cosmology. By considering both the time-delay distance and the angular diameter distance per system, lensing distances become more competitive to other distance probes than it has been previously appreciated. Specifically, strong lensing is useful because it does not require any cosmological assumptions except for the general theory of relativity on cosmological scales, and is independent of other distance measurements. However, instead of relying on cosmological assumptions, individual properties of the lens (such as mass distribution and the potential estimate via kinematics) should be constrained via dedicated observations of each system. This requires deep imaging and spectroscopic observations with high spatial resolution, as well as high cadence monitoring to measure the time delay. In this chapter, I outline the current status and the future perspective in strong lensing cosmography. First I review the past and current programs that discovered known time-delay lenses from imaging surveys, and list other algorithms that have been proposed to improve the performance of the current methods. Then I discuss ongoing and upcoming cadenced imaging surveys. Currently the main targets are SNe, but with deeper observations, wide-field cadenced surveys are ideal for detecting lensed transient. I introduce a proposed algorithm that will enable the automated detection of lensed SNe type Ia from cadenced surveys. Finally, I introduce follow-up observations that are required to obtain crucial data for cosmography.

5.1 Time-delay Lens Detections in Wide-field Surveys

Starting from the discovery of the twin quasar SBS 0957+561 in 1979, many of the lensed quasars are discovered through archival searches of known quasar catalogs. Wide-field surveys have been providing a wealthy pool of quasars from which the lensed quasars can be discovered. The Cosmic Lens All-Sky Survey (CLASS; [75, 17]) combined with the

Jordell Bank VLA Astrometric Survey (JVAS) focused on finding lensed radio sources with a few arcsecond of separation as lens candidates, mainly using Very Large Array (VLA) at 8.4 GHz. Candidates from these observations were then followed up by MERLIN and Very Large Baseline Array (VLBA) for better resolution to do the surface brightness and morphology test on the candidates. JVAS/CLASS have found 22 lenses, among many of which are still being studied. The SDSS Quasar Lens Surveys (SQLS; [79, 80, 81, 45, 46, 47]) first spectroscopically selected quasars in the SDSS sample, then searched for objects with extended morphology and nearby companions with the same color among the pre-selected quasars. The SQLS final catalog from the Data Release 7 (DR7) contains 62 lensed quasars that are confirmed by additional observations. The same algorithm was applied to the SDSS-III BOSS data (BOSS Quasar Lens Survey, BQLS), reporting detections of 13 spectroscopically confirmed doubly lensed quasars, and 11 apparent quasar pairs that can potentially be lensed [73]. The STRong lensing Insight in the Dark Energy Survey (STRIDES; [103]) program is using observations from the Dark Energy Survey (DES), reporting detections of two strongly lensed quasars [7]. The detections of lensed quasars are done in multiple steps of observations, as large surveys can provide only limited information about individual targets: The candidate confirmation process is described in section 5.2.1. In section 5.1.1, I will discuss various approaches that are designed to be applied to large data to search for the lens candidates.

5.1.1 Finding Lens Candidates by Image Inspection

Upcoming large surveys promise discoveries of hundreds of thousands of strong lenses, among which hundreds will be time-delay lenses. However, from wide-field, deep survey volumes, a problem of finding strongly lensed systems still remains. In searching for lenses from imaging data, several image inspection algorithms have been proposed. The key requirements are speed, high completeness and purity; these algorithms are tested with both real and simulated lenses and non-lenses, to give estimates of their performances.

CHITAH, an algorithm presented in [20], uses simple lens mass models to test the image configuration. The method uses color information to disentangle the lens light from the image light, as source quasars tend to be bluer than the lens galaxies, which are typically elliptical galaxies. The algorithm then tests whether the images can be mapped back to a single source position under the given simple mass model. The method can detect both double and quadruple lenses. However, the true positive rate is shown to be higher in detecting quadruple lenses, which are more useful in cosmography. The algorithm was tested using a set of simulated lenses that resembles the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS). For the best subset, which are bright ($m_z < 22.5$) quadruply lensed images with wide separations ($r_{\rm ein} > 1.1''$), the true positive rate was $\sim$90% and the false positive rate was $\sim$3%. The method has been applied to Hyper-Suprime Cam (HSC) data, cross-matched with the Galaxy And Mass Assembly (GAMA) survey [21], finding 10 galaxy-galaxy lensing candidates.

[6] proposed the use of data mining to find lensed quasars using imaging data. The algorithm takes two steps: the first uses an artificial neural network to preselect *targets*

using their apparent magnitude in each color band and the second moment, which are position angles and axis ratios of the galaxies in the images. Then a machine learning algorithm is applied to the targets, trained using color-morphology information on a pixel-by-pixel basis, to confirm the final *candidates*. It has been claimed that this method can be easily extended and its performance will be improved by the use of temporal information from time domain surveys. The strength of the method is in its speed, while also achieving 60% completeness and 70% purity. By the nature of machine learning, the performance will be significantly improved with bigger datasets. Also, with tremendous ongoing advances in machine learning, this method has big potential in automated candidate selection for upcoming wide-field surveys.

Space Warps [67] is a program proposed to use volunteers from the general public to crowd-source visual inspections. They conducted a test using samples from CFHTLS and using its mock simulations. They provided color-composite lensed images to the volunteers where the images are a mixture of observed and simulated lenses. The participants are asked to separate lensed objects into three categories: lensed galaxies, lensed quasars and cluster lenses. Each time the participants make choices, they get feedback in real time to improve their performances. From the test, the program reported 90% completeness and 30% purity of the final candidates.
These methods can be used together to complement each other, to provide reliable candidates for large volume surveys which should be followed-up for the purpose of cosmography. The details of follow-ups are descried in section 5.2.
As time-delay lenses have variable sources, surveys that are designed to find transients also suits to find these objects. In the following section, I discuss the status of present and future cadenced surveys and introduce a proposed strategy to find lensed SNe type Ia in these surveys.

5.1.2 Cadenced Wide-Field imaging Surveys

Visiting the same region in the sky periodically, cadenced surveys have opened a new era of time-domain cosmology. Transient sources, such as SNe and quasars, can be easily found by comparing the brightness on the sky observed at different times.
Ongoing cadenced surveys, such as High Cadence Transient Survey (HiTs; [35]) and Intermediate Palomar Transient Factory (iPTF; [60]) are mainly aiming for discoveries of SNe in the local Universe. However, with deeper observations, finding lensed transients is also possible.
LSST is the most ambitious upcoming cadenced wide-filed survey planned to date. One of the science goals of LSST [48] is to discover rare transient events in the optical sky, enabling detections of lensed quasars and SNe. With six bands covering wavelengths from $320 - 1050$nm, The survey is designed to have universal cadence (in the optimistic case, 3-4 days) and will uniformly cover 18000 square degrees of the sky in 10 years of its main deep-wide-fast survey mode. [78] has predicted that LSST will be able to find around 8000 lensed quasars and 130 lensed SNe.

5.1.3 Lensed SNe type Ia in Cadenced Surveys

In the case of lensed type Ia SNe, automated detection is easier due to their known brightness. The first lensed type Ia SNe was detected by iPTF in 2016 [38]. The target can be found among over a billion detections of variability, first narrowed down to a few hundred based on an algorithm presented in [37]. The method relies on three pieces of information: i) the absolute brightness of a type Ia SNe at its peak luminosity is roughly known, ii) unresolved and magnified, multiply imaged lensed type Ia SNe will appear brighter than the unlensed type Ia SNe and iii) early type galaxies only host type Ia SNe, while other types of galaxies can host both type Ia SNe and core collapse SNe. Due to the sharp 4000 Angstrom break in the quiescent early type galaxies, the photometric redshift of the host galaxies can be easily determined. After assuming a cosmology, the luminosity distance, and thus the apparent brightness of the SNe can be determined. Thus, from ii), any object that appears brighter than the predicted apparent magnitude at the host galaxy redshift becomes lensed SNe candidate. The algorithm suggests cross-matching existing samples of galaxies with secure photometric redshifts against variability detections from cadenced surveys (e.g. SDSS and iPTF), so that the variables hosted by the quiescent galaxies can be selected. Then the absolute brightness cut is applied to the sample to finalize the candidate selection. AGNs or core-collapse SNe are potential contaminants of the method, but these can still be used for the purpose of lensing cosmography.

5.2 Follow-ups

For individual lens candidates, high-resolution imaging and spectroscopic follow-ups are both required, to confirm the candidacy, to construct the lens mass model, to constrain the lens environment and its line-of-sight structures and to obtain kinematics information of the lens. High cadenced monitoring is also required to measure the time delay. Also, spatially resolved kinematics data is becoming more important, as it can be used to break the mass-anisotropy degeneracy and improve the constraints on angular diameter distance.

5.2.1 Candidate Confirmation

As the purity of the candidates from the binary classification test on imaging data is not 100%, a follow-up observation is required to confirm individual systems as lenses. SQLS proposed strategic use of optical imaging, spectroscopy and near-infrared imaging to confirm the candidates [45, 46]. When the candidates were selected morphologically, optical imaging can reject a large fraction of non-lenses. When the candidates were selected based on their color, the false-positive rate was higher than that of morphological selection, and spectroscopic observation of stellar components and the imaging of the lensing object for individual candidates are both required to confirm gravitational lens systems [47].

5.2.2 High-resolution Imaging: Adaptive Optics (AO)

With limited resources in high-resolution imaging, such as HST, AO is becoming a prominent tool in improving the spatial resolution of ground-based imaging observations. The main idea of AO is to measure the deformation of the incoming wavefront in real time using guide stars (sometimes with a laser), and change the surface of the mirror accordingly such that the deflected wavefront will be corrected. This technique is applied to instruments in observatories such as Keck, and have been shown to improve the resolution from $\sim 1''$ to $0.03 - 0.06''$. However, modeling lensed quasars with AO observations is a challenge as the PSF is unknown and unstable. The light from the central point source component (AGN) is hard to separate from the extended host galaxy, unlike non-AO observations where PSF is simpler and for most purposes can be approximated as a Gaussian. Thus AGN light acts as a contaminant to the image arcs, preventing precise modeling of the lens mass using the extended components of the images. To overcome the limitation, an iterative reconstruction of the point function (PSF) has been carried out for a lensed system, RXJ1131-1231 and two mock simulations of it in [22]. Their analysis simultaneously models the PSF, the point source (AGN) and the extended component (host galaxy) of the images as well as the lens light. They have successfully separated these three components, showing that the Keck telescope AO can provide spatial resolution twice as high as HST. They also have shown that the resulting cosmological parameters constrained by the time-delay distance of the system with AO observations are consistent with those from HST within 1-σ. Thus, further ground-based high-resolution imaging observations with AO will be a good alternative in following up lens candidates to allow further modeling of lenses with good precision, with the eventual goal of a few per cent level constraint on the Hubble constant.

5.2.3 Monitoring the Time delay

The first time-delay measurement of a lensed quasar, QSO0957+561A was made in 1984 [34]. With only a few measurements per year, however, the resulting time-delay remained controversial with uncertainty of half a year, and the controversy lasted until 30 years of further observation finally improved the measurement. Even in recent times, the most time-consuming component in observing time-delay lenses for cosmography is to monitor the time delay of each quasar image, which takes $\sim$ 10 years per system. Optical monitoring of known quasar lenses has been conducted by the COSmological MOnitoring of GRAvItational Lenses (COSMOGRAIL collaboration; [29]). As the cadence is the most important factor in assessing the light curve of stochastic variables, COSMOGRAIL used multiple small, but dedicated telescopes to monitor the lenses as regularly as possible. The typical frequency of visits is once per a few days during the observing season; a season lasts for $\sim$8 months and the monitoring last for $\sim$5-10 years per system. The program provided time-delay measurements for about a dozen quasar lenses to date, with precision of a few per cent.

After measuring the light curve, obtaining the time-delay between image pairs still requires thorough analysis. Time-delay Challenge (TDC; [27, 62]) did a performance com-

parison between multiple algorithms that calculate the time delay from thousands of mock light curves. The simulated light curves were created including microlensing and stochastic AGN light curves using a damped random walk. To mimic realistic observations, photometric and systematic errors are also simulated. Observing frequency and duration, i.e. cadence, season length (the continuous monitoring period per system with given cadence: typically ∼months) and campaign length (total duration of time the system is monitored: ∼ years) were controlled to match those of COSMOGRAIL and optimistic / realistic cases of LSST. Their results have shown that only 20-40% of the simulated light curves yield time-delay measurements, and the success rate depends on the season and campaign length. They also reported that the accuracy of the time delay measurement depends on the duration of the monitoring, while the precision can be improved with higher cadence. The algorithms with the best performance could achieve sub-percent accuracy, which is better than the current few percent level time-delay measurements.
On the other hand, SNe do not require similarly long follow-ups to measure time delay, as the shape of its light curve is well defined, and also the brightness decays within a few weeks. However, due to this decay in brightness, the monitoring should be done in a short period of time as the peak of the light curve is crucial in determining the characteristics of the SNe.

5.2.4 High-resolution spectroscopy

Kinematic measurements of lenses are becoming more important as the final pieces of information in time-delay lenses, as it is crucial in constraining the angular diameter distance [50], and also it can break the mass-sheet degeneracy when lens mass profile is known [114]. Specifically, it is shown that when the external convergence is constrained via simulations, spatially resolved kinematics of the lens galaxy can also breaks the mass-anisotropy degeneracy (Anowar, Treu and Agnello, 2017, submitted).
Spatially resolved kinematics of the lens requires observations with integral field unit spectrographs (IFU), which measure spectra in two dimensions. To guarantee high signal-to-noise ratio over multiple pixels in the image, the telescope should have large aperture.
Currently instruments such as OH-Suppressing Infrared Integral Field Spectrograph (OSIRIS; [59]) on Keck, and the Spectrograph for INtegral Field Observations in the Near Infrared (SINFONI; [30]) on the Very Large Telescope (VLT) are IFU instruments operating with AO, providing diffraction-limited spatial resolution. In the future, Near Infrared Spectrograph (NIRSPEC) on James-Webb Space Telescope (JWST) and Infrared Imaging Spectrometer (IRIS; [112]) on the Thirty-Meter Telescope (TMT) will be available for studying spatially resolved lens kinematics as well.

5.3 Discussion & Conclusion

Strong lensing cosmography is a relatively new tool in cosmology. While the idea was proposed a few decades ago, the observational data only became available in the past

ten years. In the future, however, as large quantities of high-quality data are becoming available, lensing cosmography will be a prominent tool in improving our understanding of the Universe. As I reviewed in this chapter, various works have been put together to expand the number of known time-delay lenses and to overcome the observational challenges, both in terms of hardware and algorithms. On the cosmological side, as stated in chapter 1 of this thesis, H0LiCOW will provide the baseline for measuring time-delay distances. Their data can also be used for the angular diameter distance measurements that are presented in this thesis. The combination of these two distances per lens will open a new era for strong lensing cosmography; as the number of well-measured time-delay lenses increases, lensing cosmography can easily reach the statistical and systematic uncertainties assumed in [51]. This will provide a unique opportunity to probe the expansion history of the low-redshift Universe.

Appendix A

Deflection angle of an arbitrary power-law density profile

We derive the expression for a deflection angle near a galaxy with the density profile following a power-law with arbitrary density profile index. When the density profile is given as equation (2.5), the mass contained within a radius r is

$$M(r) = \int_0^r 4\pi r^2 \rho_0 r_0^{\gamma'} r^{-\gamma'} dr = \frac{4\pi\rho_0 r_0^{\gamma'}}{-\gamma'+3} r^{-\gamma'+3}, \tag{A.1}$$

which yields an acceleration given by

$$\vec{g}(\vec{r}) = -\frac{4\pi G\rho_0 r_0^{\gamma'}}{3-\gamma'} r^{-\gamma'}\vec{r}, \tag{A.2}$$

on the test mass located at the radius r. According to the post-Newtonian approximation in General Relativity, the rate of change of the direction of the velocity vector of the test mass, $\vec{u}$, is given as

$$c^2 \frac{d\vec{u}}{dt} = -2\vec{u} \times (\vec{u} \times \vec{g}). \tag{A.3}$$

We define a new parameter α to be the angle by which the light is deflected as it passes near the lens galaxy. In the cases we consider, the deflection angle will be small. Thus, we can choose coordinates such that the path of the light is roughly along the x-axis, and the line connecting the center of the lens galaxy to the point of the closest approach is along the y axis. Again, because the deflection angle is small, we use the thin lens approximation, namely, light is bent sharply at the closest approach to the lens. Thus, the separation from the center of lens to the light path, r, becomes $r^2 = b^2 + x^2$, and, more importantly, $\vec{u} \times (\vec{u} \times \vec{r}) = -c^2\vec{b}$.

We define the deflection angle at the lens plane, $\hat{\alpha}$, as the total change in the photon propagation direction, and the magnitude of the deflection angle as α. Then,

$$\hat{\alpha} \equiv \frac{1}{c}\int d\vec{u} = -\alpha\frac{\vec{r}}{r}, \tag{A.4}$$

where the minus sign indicates that the deflection happens toward the lens center. Then α becomes

$$\begin{aligned}\alpha &= \frac{8\pi G\rho_0 r_0^{\gamma'}}{c^2(3-\gamma')}\int_{-\infty}^{\infty} br^{-\gamma'}dx \\ &= \frac{8\pi G\rho_0 r_0^{\gamma'}}{c^2(3-\gamma')}b\int_{-\infty}^{\infty}\frac{dx}{(x^2+b^2)^{\gamma'/2}} \\ &= \frac{8\pi G\rho_0 r_0^{\gamma'}b^{2-\gamma'}}{c^2(3-\gamma')}\frac{\sqrt{\pi}\;\Gamma[\frac{1}{2}(-1+\gamma')]}{\Gamma(\frac{\gamma'}{2})},\end{aligned} \tag{A.5}$$

assuming $\gamma' > 1$.

Appendix B

Lensing constraints on H_0

B.1 H_0 in owCDM and ow_zCDM models

We show the constraints on H_0 from D_A and $D_{\Delta t}$ combined with the Planck distance prior. Figure B.1 shows the expected 1-σ uncertainties in h from strong lenses combined with the Planck distance priors. As $D_{\Delta t}$ is mostly sensitive to H_0, the constraining power of $D_{\Delta t}$ + Planck (red dotted line) is more powerful than that of D_A + Planck (blue dashed line). When w is fixed as a constant (owCDM model, left panel), $D_{\Delta t}$ + Planck are more powerful than $D_{\Delta t}$ + D_A (black dot-dashed line). When w is allowed to vary (ow_zCDM model, right panel), however, $D_{\Delta t}$ + D_A is more powerful than $D_{\Delta t}$ + Planck. This is due to the degeneracies between H_0, Ω_k and w from the linear CMB constraints alone [28], which cannot be broken by $D_{\Delta t}$. However, ref. [43] has shown that the main degeneracy from CMB constraints is between w and H_0, and as shown in section 4.2.1, the combination of lensing distances is powerful in breaking the degeneracy between Ω_k and w. Thus, the combination of Planck and the lensing distances shows 30% improvement in constraining h.

B.2 H_0 in flat ΛCDM model

We show the constraints on H_0 for the ΛCDM model in figure B.2. Assuming that 5% precision measurements in individual distance (both D_A and $D_{\Delta t}$) are achievable from lens systems, 10 lenses are enough to measure the Hubble constant to the same precision as Planck. The number of required lenses to achieve the same precision increases to 25 if constraints are from $D_{\Delta t}$ only.

B.3 Constraints assuming the flat universe

In section 4.2.1, we have shown that the lensing distances are powerful probes for the curvature of the universe. Specifically, D_A and $D_{\Delta t}$ respond to curvature differently in

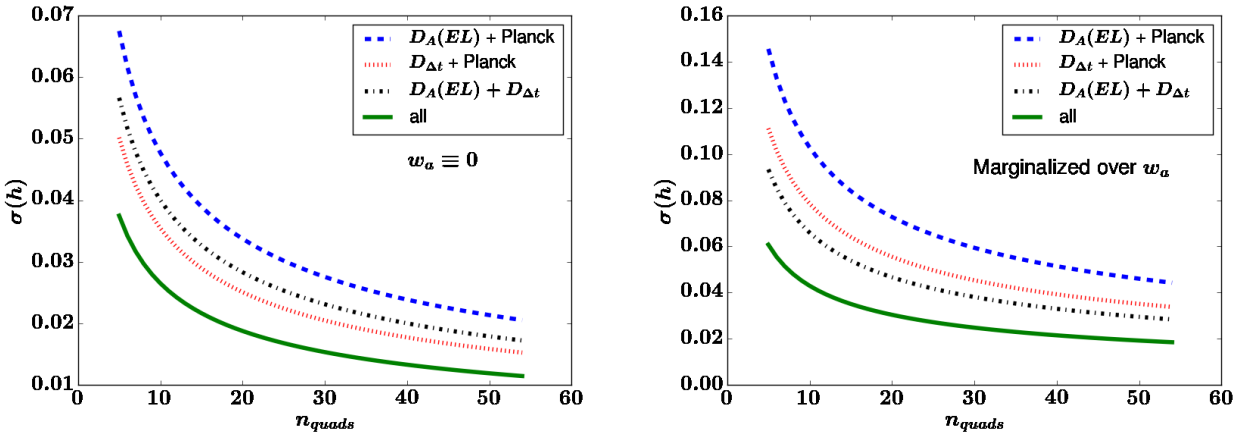

Figure B.1: The 1-σ uncertainty in $h = H_0/100$ km/s/Mpc from time-delay lenses as a function of the number of quadruply imaged lenses for the (left) owCDM model, and (right) ow_zCDM model.

w_0-w_a plane, thus the combination of two gives a strong constraint on Ω_k. We repeat the same analysis for the flat universe model ($\Omega_k \equiv 0$). The model parameters are summarized as

$$\vec{\theta} \in \{\Omega_m, w, h\} \quad \text{(flat } w\text{CDM model)}, \tag{B.1}$$

and

$$\vec{\theta} \in \{\Omega_m, w_0, w_a, h\} \quad \text{(flat } w_z\text{CDM model)}. \tag{B.2}$$

The constraining contours for these models are shown in figures B.3 and B.4.

Figure B.3 shows that under the flatness assumption, the constraints from $D_{\Delta t}$ + Planck are already as tight as those from $D_A + D_{\Delta t}$ + Planck, i.e., the constraining power from D_A in flat universe is minor. Figure B.4 shows that the 55 lenses combined with Planck still constrain the equation of state better as compared to Planck + JLA and Planck + BOSS for the flat wCDM model (left panel), and comparably well as Planck + BOSS for the flat w_zCDM model (right panel).

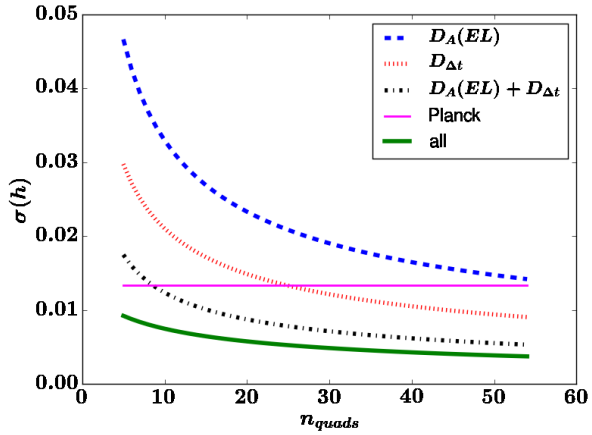

Figure B.2: Same as figure B.1, but for the flat ΛCDM model. Here the blue dashed, red dotted and black dash-dot lines are from the lensing distances alone, not combined with Planck. We show the constraints from Planck as the horizontal magenta solid line and Planck + lensing distances as the green solid line. Planck-precision constraint in h is achievable with 10 lenses when we use both D_A and $D_{\Delta t}$, while we need 25 lenses to achieve the same constraint from $D_{\Delta t}$ alone.

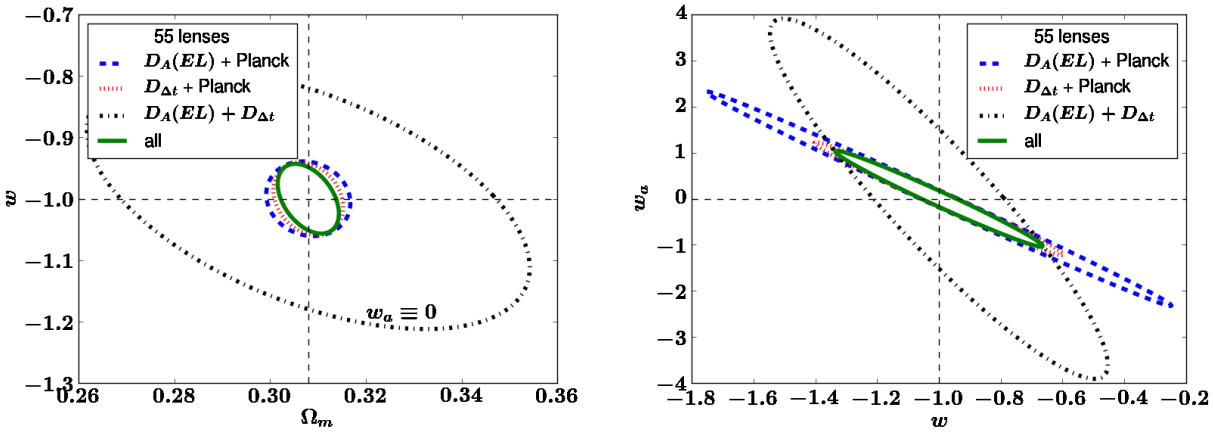

Figure B.3: Same as figure 4.5, but for the flat (left) wCDM and (right) w_zCDM model.

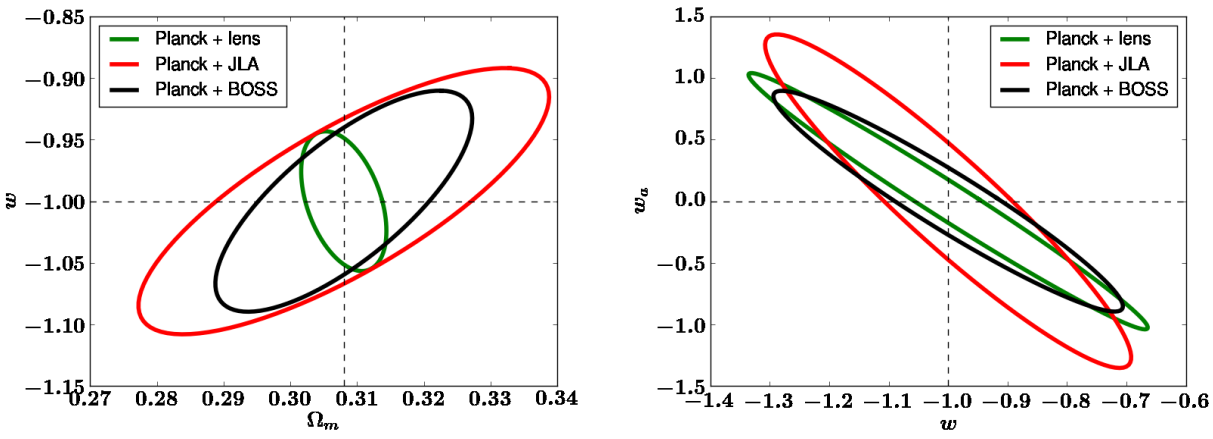

Figure B.4: Same as figure 4.6, but for the flat (left) wCDM and (right) w_zCDM model.

Appendix C

Full constraints and the Hubble Diagram from Lensing Angular Diameter Distance combined with SNe type Ia

This appendix show the full cosmological and nuisance parameter constraints from angular diameter distances to two lenses and 740 SNe from JLA sample. The resulting cosmological and nuisance parameters are summarized in tables 3.1 and 3.2, respectively.

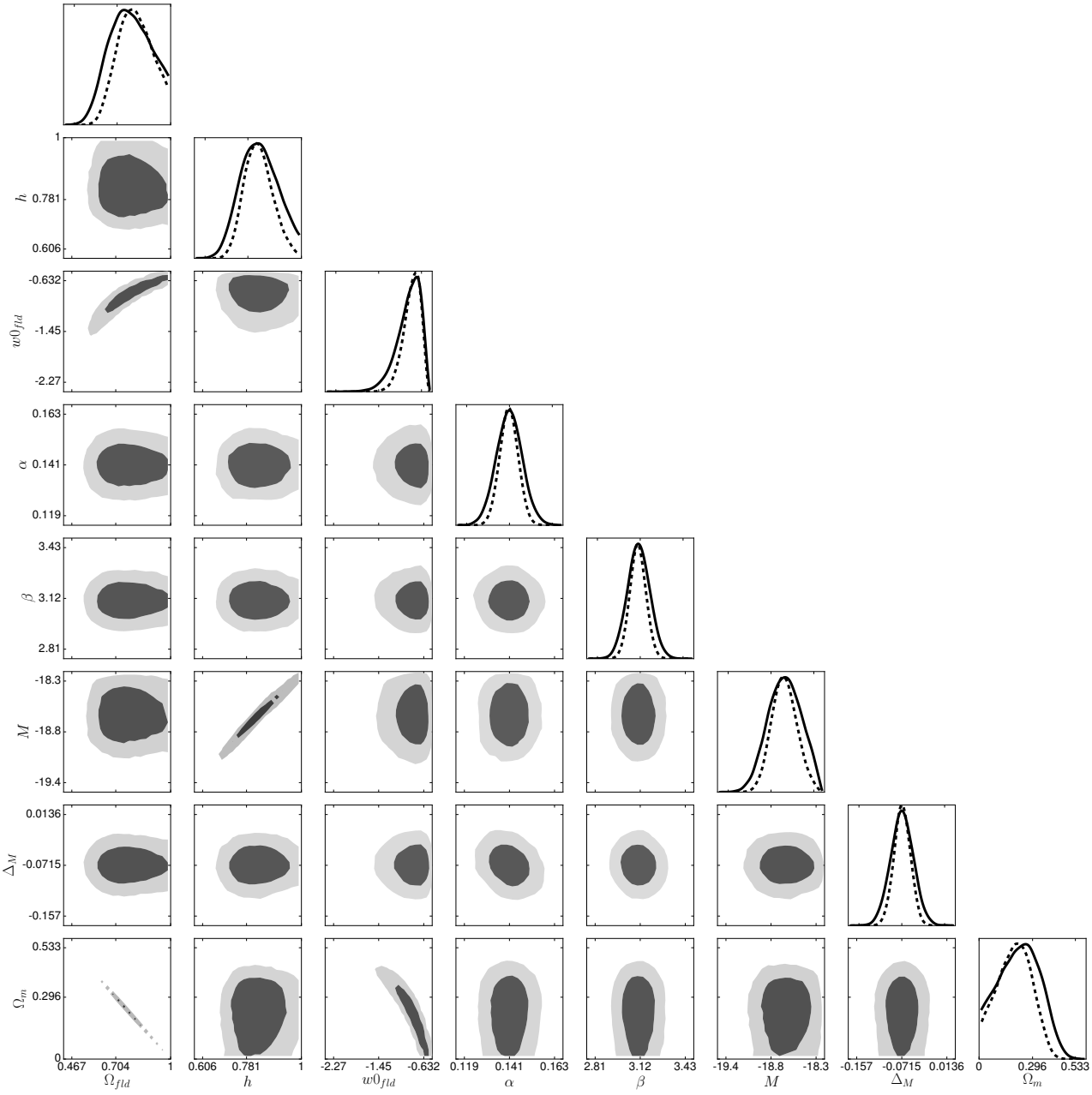

Figure C.1: Same as figure 3.8, but for flat wCDM model.

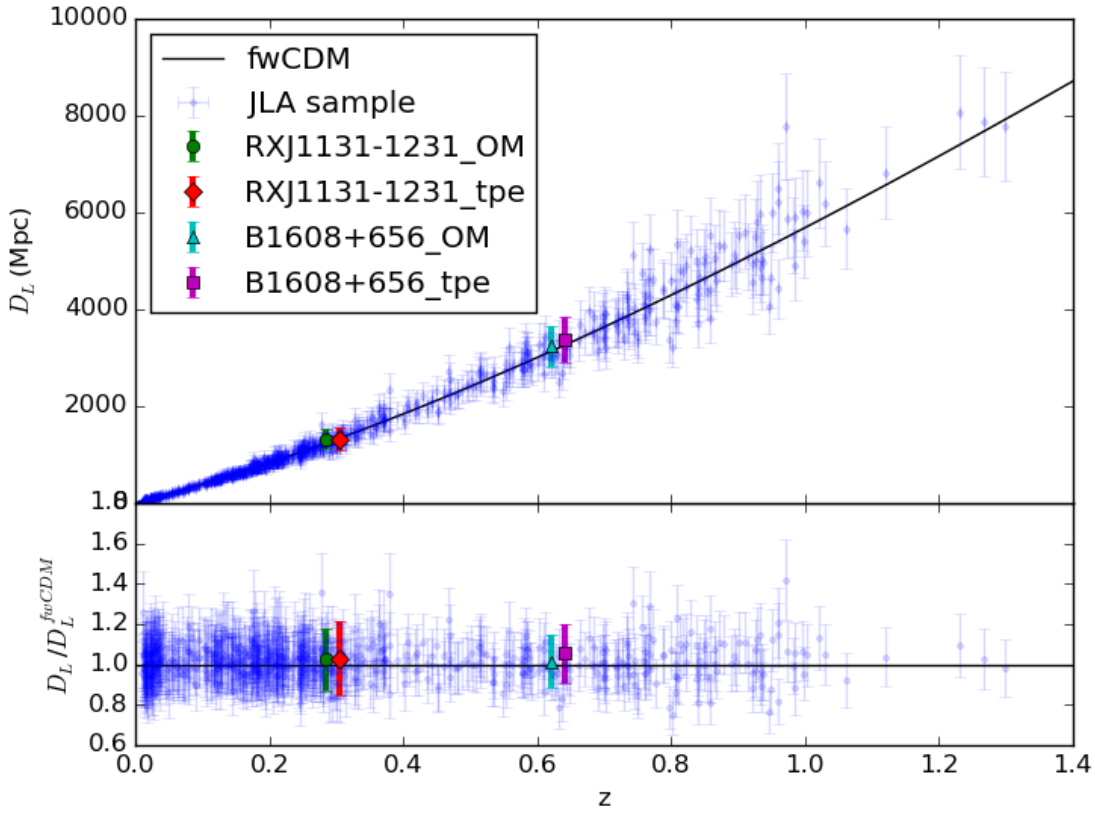

Figure C.2: Same as figure 3.9, but for flat wCDM model fitted with the cosmological and nuisance parameters constrained from figure C.1.

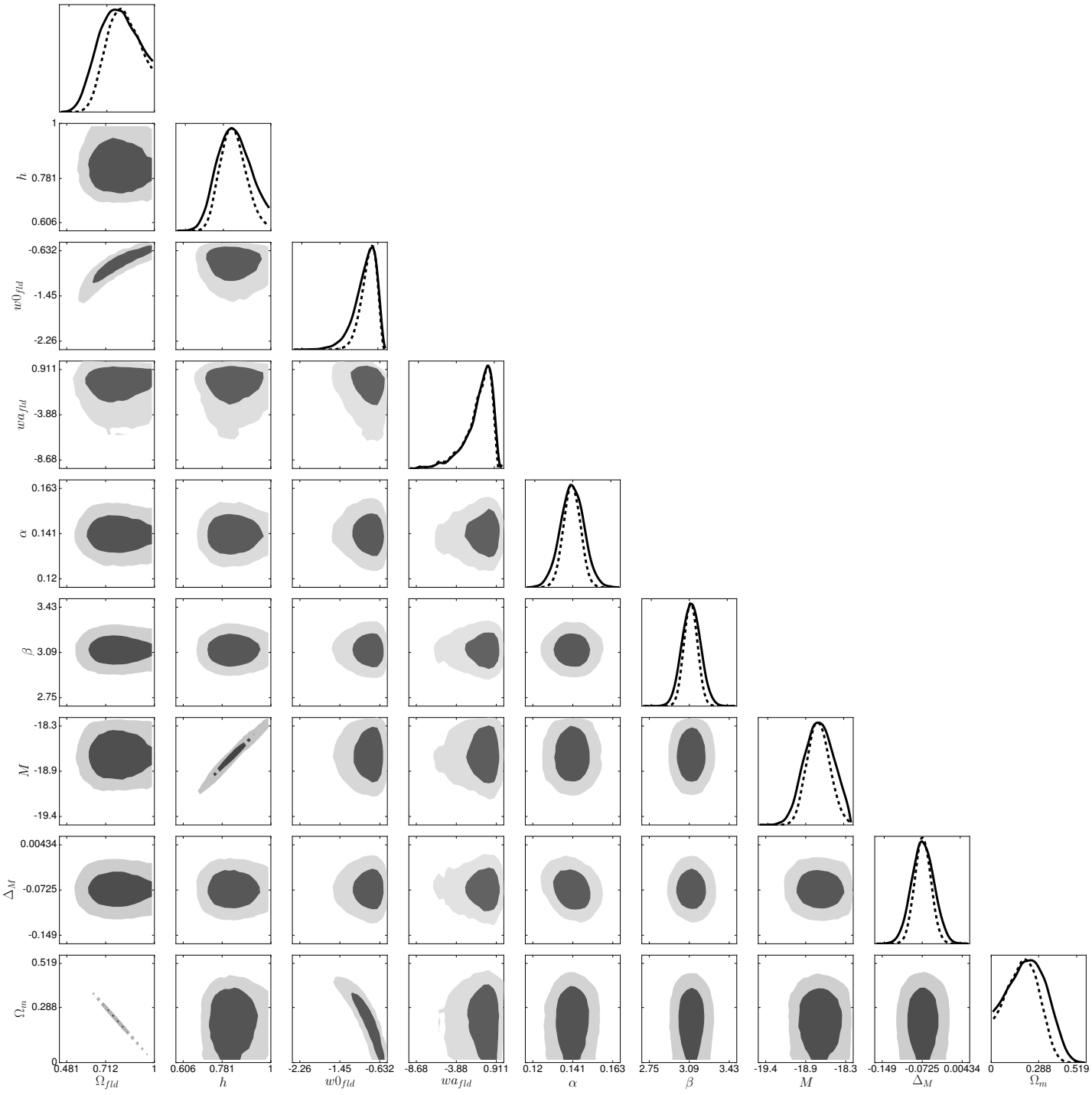

Figure C.3: Same as figure 3.8, but for flat w_aCDM model.

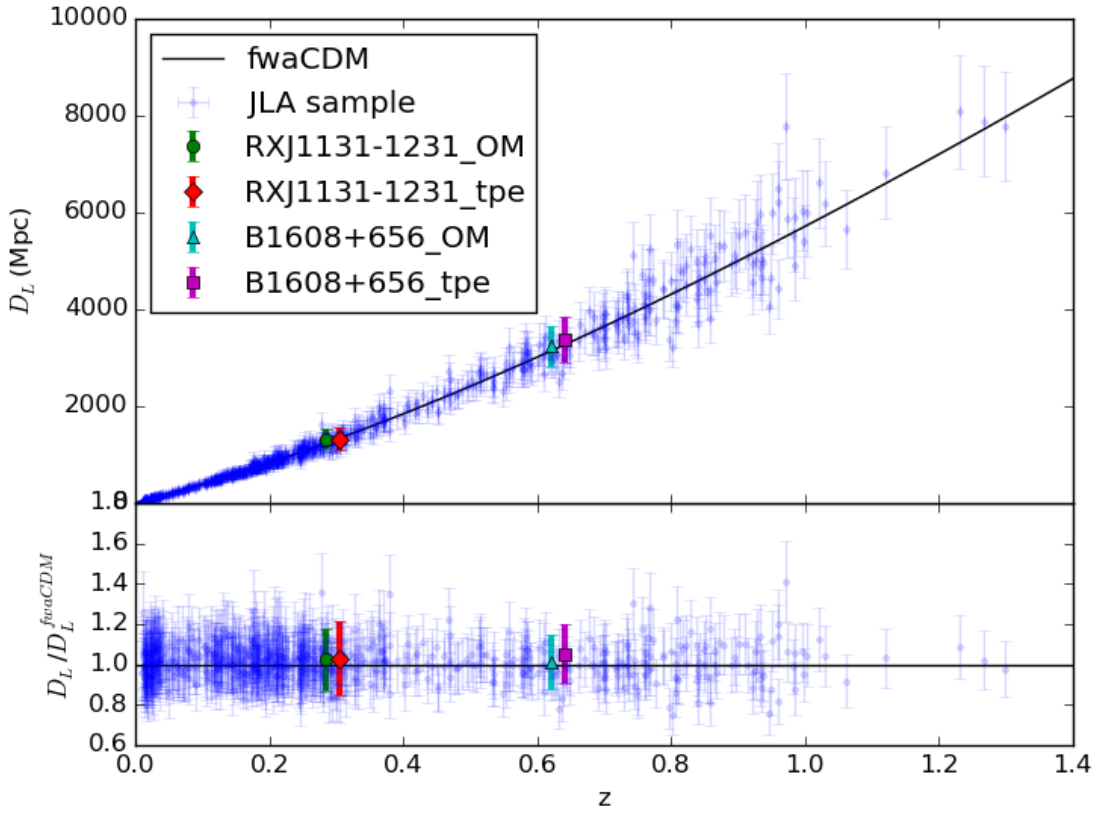

Figure C.4: Same as figure 3.9, but for flat w_aCDM model fitted with the cosmological and nuisance parameters constrained from figure C.3

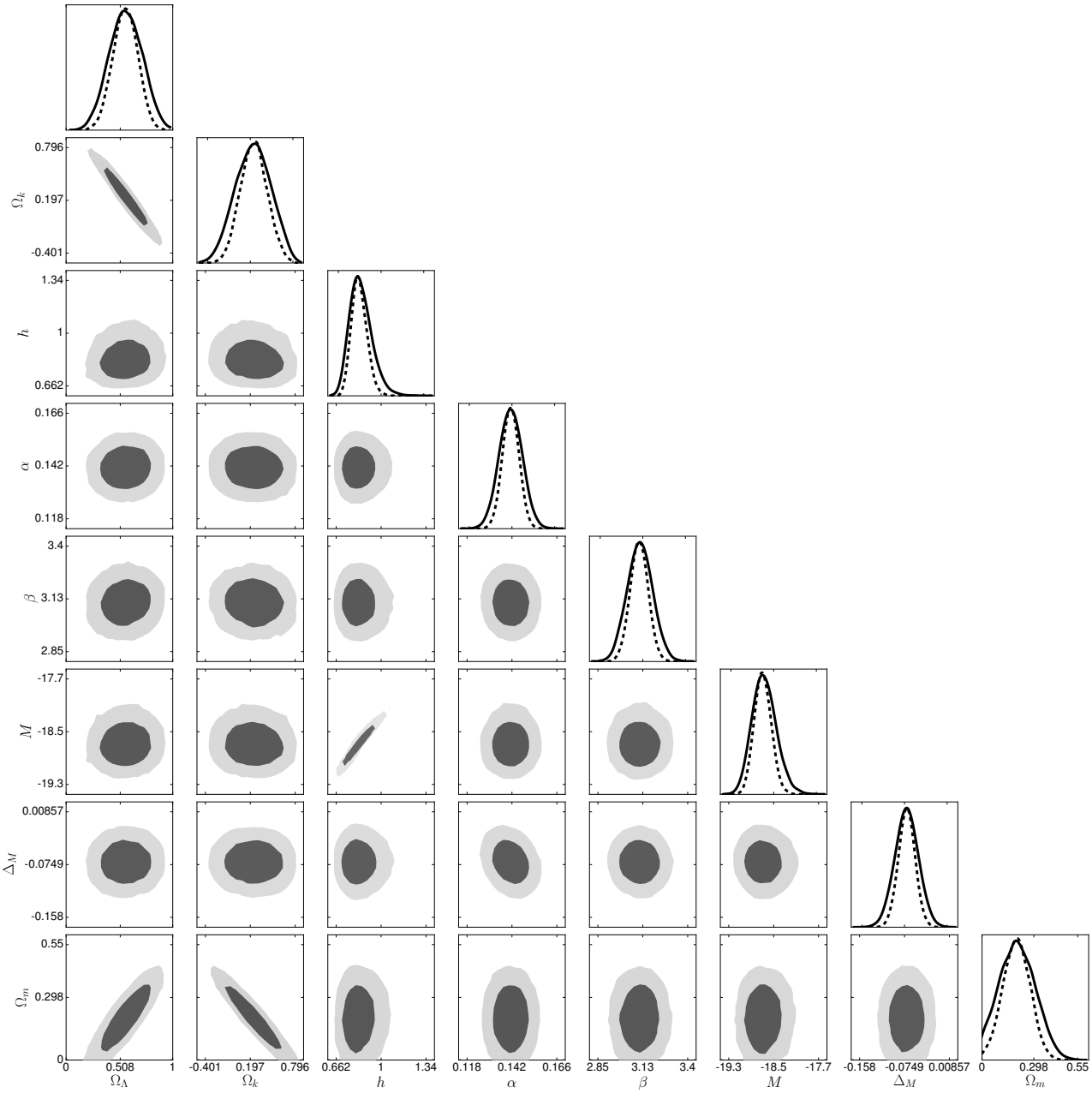

Figure C.5: Same as figure 3.8, but for open ΛCDM model.

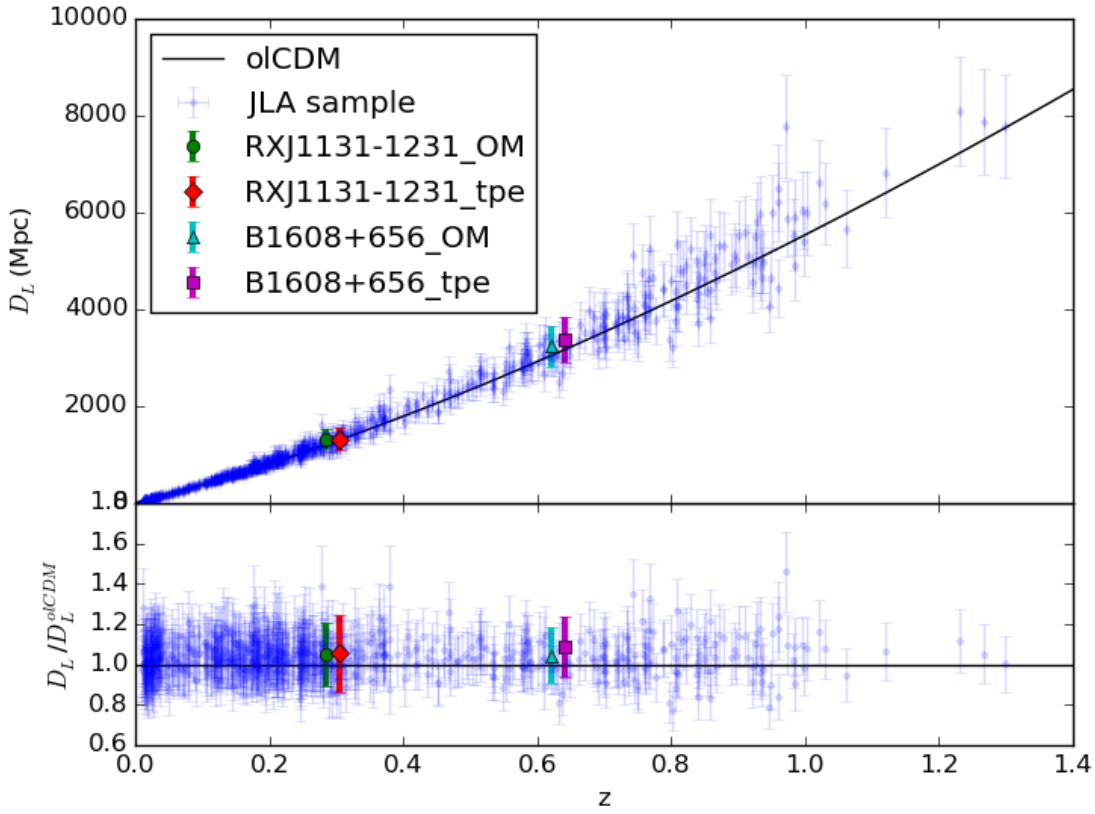

Figure C.6: Same as figure 3.9, but for open ΛCDM model fitted with the cosmological and nuisance parameters constrained from figure C.5

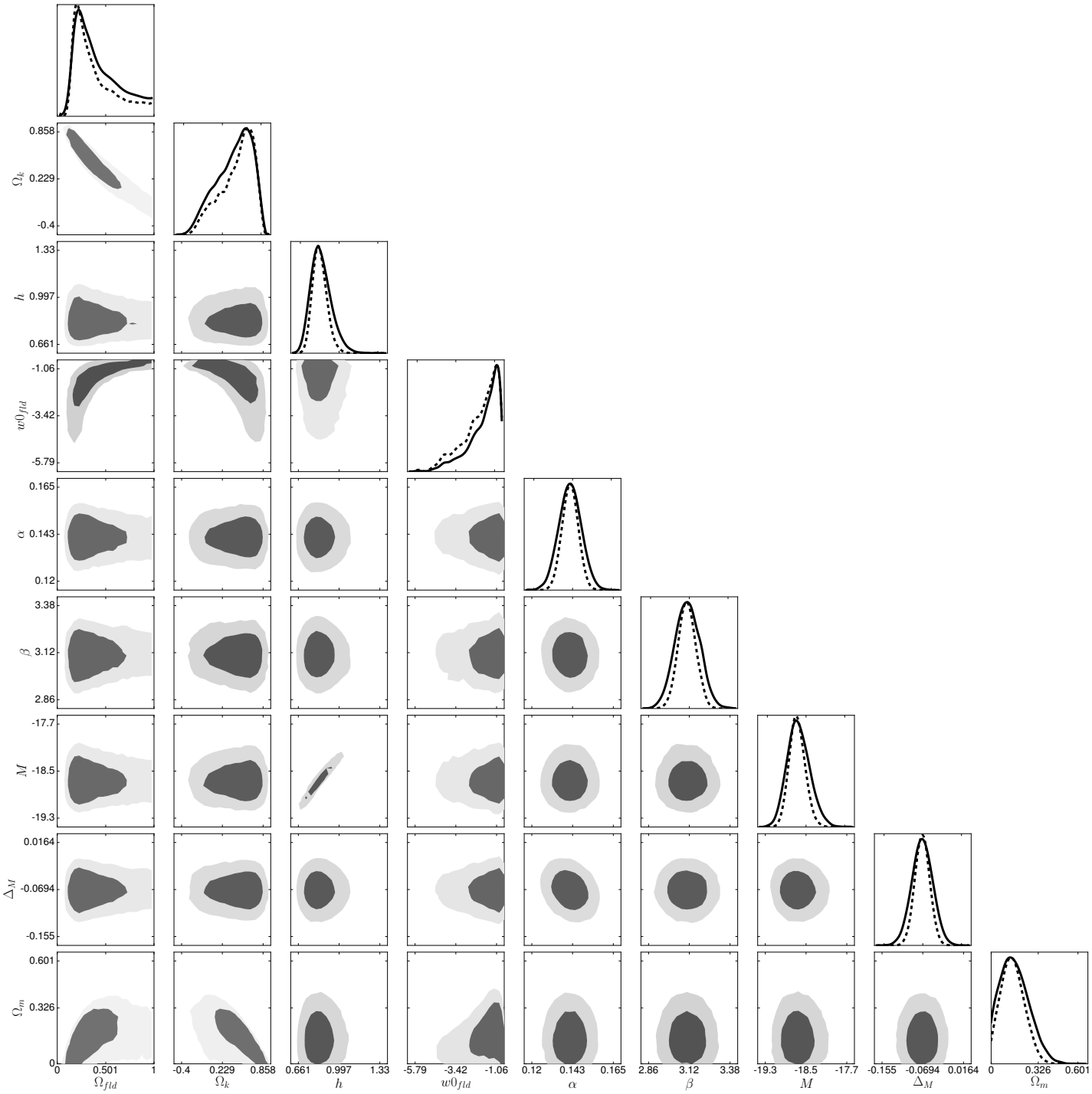

Figure C.7: Same as figure 3.8, but for open wCDM model.

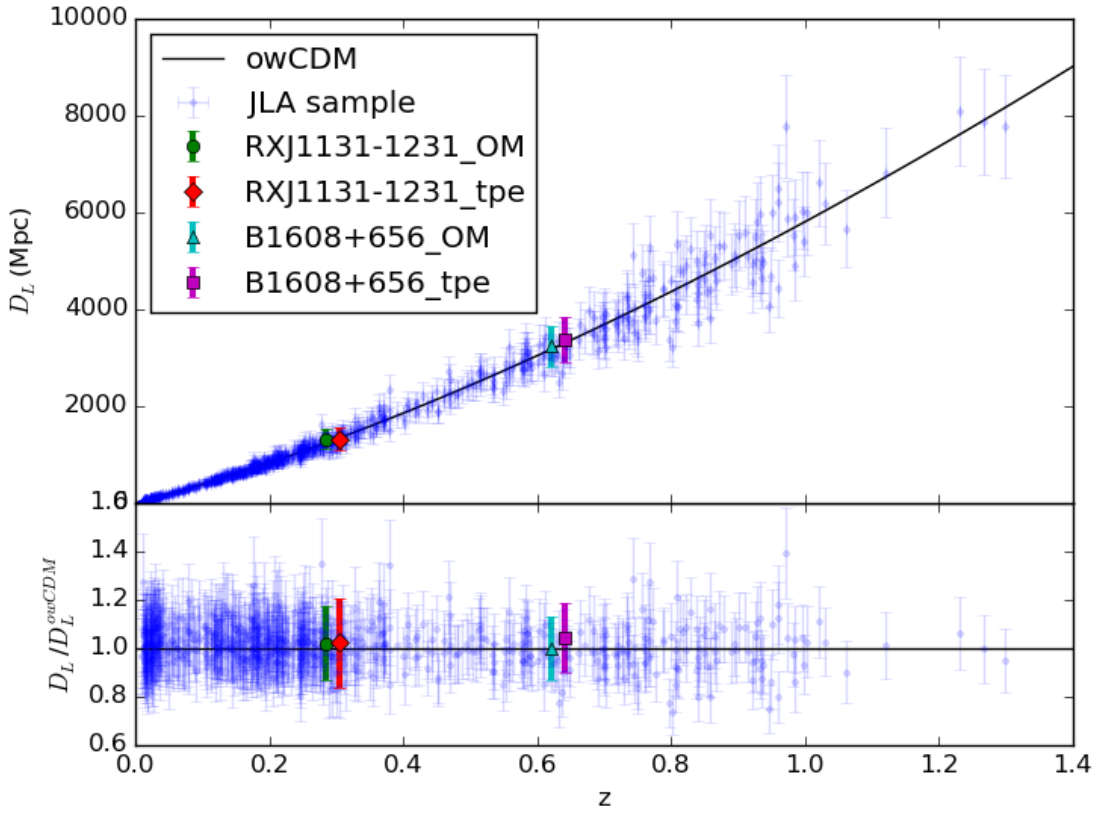

Figure C.8: Same as figure 3.9, but for open wCDM model fitted with the cosmological and nuisance parameters constrained from figure C.7.

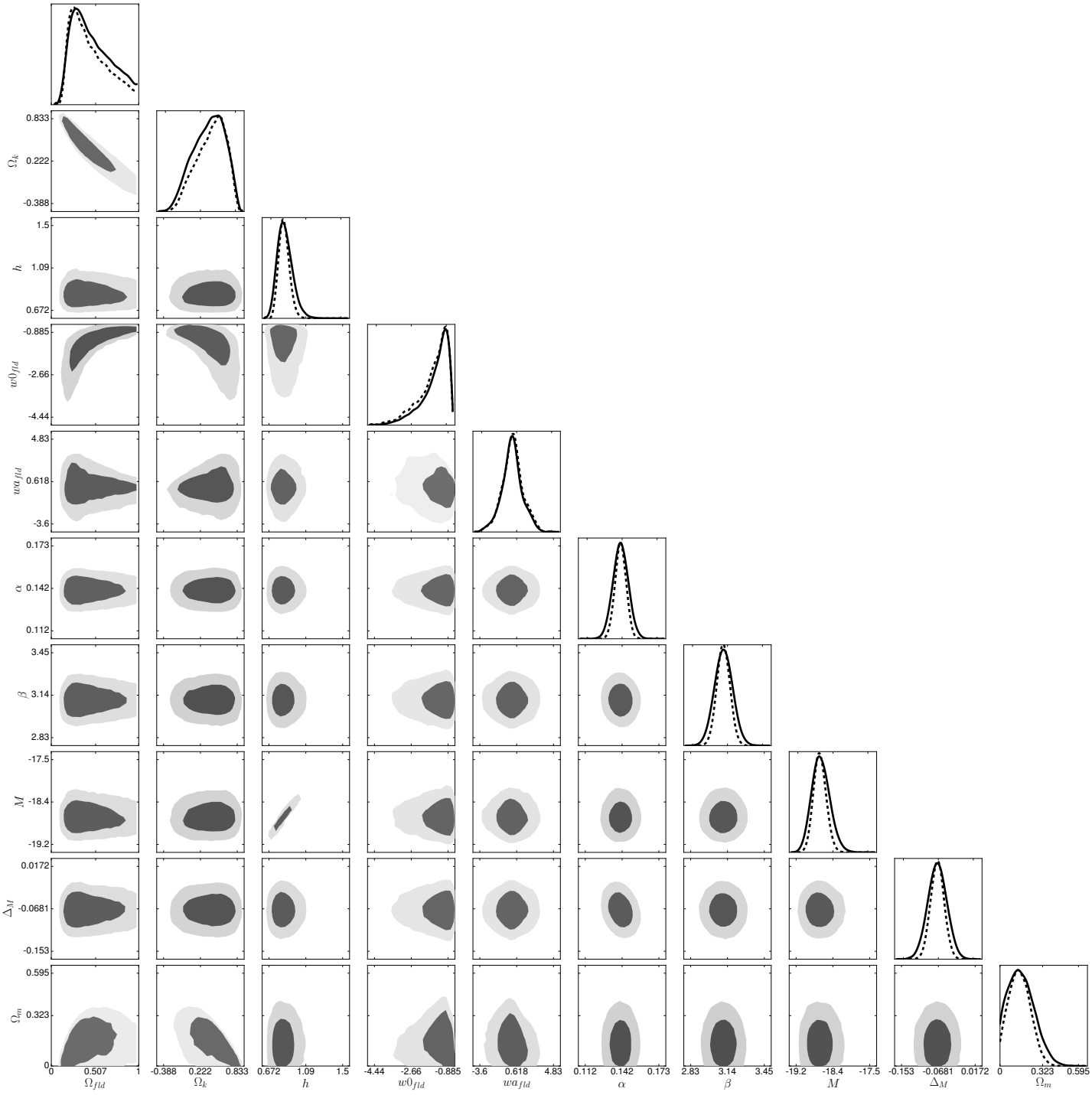

Figure C.9: Same as figure 3.8, but for open w_aCDM model.

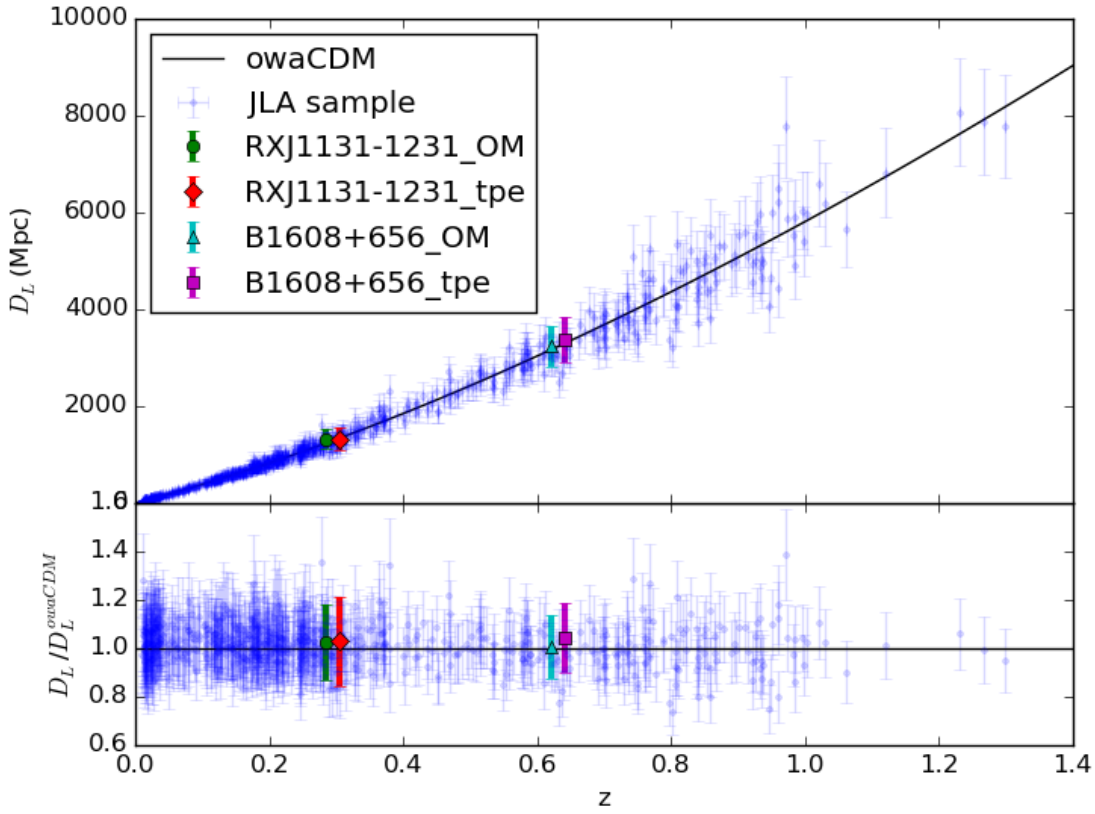

Figure C.10: Same as figure 3.9, but for open w_aCDM model fitted with the cosmological and nuisance parameters constrained from figure C.9.

Bibliography

[1] A. Abate et al. : *Large Synoptic Survey Telescope: Dark Energy Science Collaboration*, 2012.

[2] P.A. Abell et al. : *LSST Science Book, Version 2.0*, 2009.

[3] P.A.R. Ade et al. , *Astron. Astrophys.* **594** (2016), A13.

[4] P.A.R. Ade et al. , *Astron. Astrophys.* **594** (2016), A14.

[5] A. Agnello, N. Evans, A.J. Romanowsky und J. Brodie, *Mon.Not.Roy.Astron.Soc.* **442** (2014), 3299.

[6] A. Agnello, B.C. Kelly, T. Treu und P.J. Marshall, *Mon. Not. Roy. Astron. Soc.* **448** (2015), 1446.

[7] A. Agnello et al. , *Mon. Not. Roy. Astron. Soc.* **454** (2015), 1260.

[8] A. Albrecht et al. , *ArXiv e-prints* (2009).

[9] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, M. Blanton et al. , *Mon.Not.Roy.Astron.Soc.* **428** (2013), 1036.

[10] B. Audren, J. Lesgourgues, K. Benabed und S. Prunet, *JCAP* **1302** (2013), 001.

[11] M.W. Auger, T. Treu, A.S. Bolton, R. Gavazzi, L.V.E. Koopmans, P.J. Marshall, L.A. Moustakas und S. Burles, *Astrophys.J.* **724** (2010), 511.

[12] M. Barnabè, O. Czoske, L.V.E. Koopmans, T. Treu und A.S. Bolton, *Mon.Not.Roy.Astron.Soc.* **415** (2011), 2215.

[13] J.L. Bernal, L. Verde und A.G. Riess, *JCAP* **1610** (2016), 019.

[14] M. Betoule et al. , *Astron.Astrophys.* **568** (2014), A22.

[15] G.S. Bisnovatyi-Kogan und O.Yu. Tsupko, *Grav. Cosmol.* **15** (2009), 20.

[16] V. Bonvin et al. , *Mon. Not. Roy. Astron. Soc.* (2016).

[17] I.W.A. Browne et al. , *Mon. Not. Roy. Astron. Soc.* **341** (2003), 13.

[18] M. Cappellari, R. Bacon, M. Bureau, M. Damen, R.L. Davies et al. , *Mon. Not. Roy. Astron. Soc.* **366** (2006), 1126.

[19] M. Cappellari, A.J. Romanowsky, J.P. Brodie, D.A. Forbes, J. Strader, C. Foster, S.S. Kartha, N. Pastorello, V. Pota, L.R. Spitler, C. Usher und J.A. Arnold, *Astrophys. J. Lett.* (2015).

[20] J.H.H. Chan, S.H. Suyu, T. Chiueh, A. More, P.J. Marshall, J. Coupon, M. Oguri und P. Price, *Astrophys. J.* **807** (2015), 138.

[21] J.H.H. Chan et al. , *Astrophys. J.* **832** (2016), 135.

[22] G.C.F. Chen et al. , *Mon. Not. Roy. Astron. Soc.* **462** (2016), 3457.

[23] E. Churazov, S. Tremaine, W. Forman, O. Gerhard, P. Das et al. , *Mon.Not.Roy.Astron.Soc.* **404** (2010), 1165.

[24] A.J. Cuesta, L. Verde, A. Riess und R. Jimenez, *Mon. Not. Roy. Astron. Soc.* **448** (2015), 3463.

[25] N. Dalal und C.S. Kochanek, *Astrophys. J.* **572** (2002), 25.

[26] W. Dehnen, *Mon.Not.Roy.Astron.Soc.* **265** (1993), 250.

[27] G. Dobler, C. Fassnacht, T. Treu, P.J. Marshall, K. Liao, A. Hojjati, E. Linder und N. Rumbaugh, *Astrophys. J.* **799** (2015), 168.

[28] G. Efstathiou und J.R. Bond, *Mon. Not. Roy. Astron. Soc.* **304** (1999), 75.

[29] A. Eigenbrod, F. Courbin, C. Vuissoz, G. Meylan, P. Saha und S. Dye, *Astron. Astrophys.* **436** (2005), 25.

[30] F. Eisenhauer et al. , *Proc. SPIE Int. Soc. Opt. Eng.* **4841** (2003), 1548.

[31] E.E. Falco, M.V. Gorenstein und I.I. Shapiro, *Astrophys.J* **289** (1985), L1.

[32] C. Fassnacht, E. Xanthopoulos, L. Koopmans und D. Rusin, *Astrophys.J.* **581** (2002), 823.

[33] C.D. Fassnacht, D.S. Womble, G. Neugebauer, I.W.A. Browne, A.C.S. Readhead, K. Matthews und T.J. Pearson, *Astrophys.J.* **460** (1996), L103.

[34] R. Florentin-Nielsen, *Astron. Astrophys.* **138** (1984), L19.

[35] F. Frster et al. , *Astrophys. J.* **832** (2016), 155.

[36] L. Gao, J.F. Navarro, C.S. Frenk, A. Jenkins, V. Springel und S.D.M. White, *Mon. Not. Roy. Astron. Soc.* **425** (2012), 2169.

[37] D.A. Goldstein und P.E. Nugent, *Astrophys. J.* **834** (2017), L5.

[38] A. Goobar et al. , *ArXiv e-prints* (2016).

[39] M.V. Gorenstein, I.I. Shapiro und E.E. Falco, *Astrophys.J.* **327** (1988), 693.

[40] L. Hernquist, *Astrophys.J.* **356** (1990), 359.

[41] S. Hilbert, J. Hartlap, S.D.M. White und P. Schneider, *Astron. Astrophys.* **499** (2009), 31.

[42] J.W. Hsueh, L. Oldham, C. Spingola, S. Vegetti, C.D. Fassnacht, M.W. Auger, L.V.E. Koopmans, J.P. McKean und D.J. Lagattuta, *ArXiv e-prints* (2017).

[43] W. Hu, *ASP Conf.Ser.* **339** (2005), 215.

[44] D. Huterer und M.S. Turner: *Constraining the properties of dark energy. Constraining the properties of dark energy*, In *Relativistic astrophysics. Proceedings, 20th Texas Symposium, Austin, USA, December 10-15, 2000.* (2001) Seiten 297–302. [AIP Conf. Proc.586,297(2001)].

[45] N. Inada et al. , *Astron. J.* **135** (2008), 496.

[46] N. Inada et al. , *Astron. J.* **140** (2010), 403.

[47] N. Inada et al. , *Astron. J.* **143** (2012), 119.

[48] Z. Ivezic, J.A. Tyson, B. Abel, E. Acosta, R. Allsman, Y. AlSayyad, S.F. Anderson, J. Andrew, R. Angel, G. Angeli, R. Ansari, P. Antilogus, K.T. Arndt, P. Astier, E. Aubourg, T. Axelrod, D.J. Bard, J.D. Barr, A. Barrau, J.G. Bartlett, B.J. Bauman, S. Beaumont, A.C. Becker, J. Becla, C. Beldica, S. Bellavia, G. Blanc, R.D. Blandford, J.S. Bloom, J. Bogart, K. Borne, J.F. Bosch, D. Boutigny, W.N. Brandt, M.E. Brown, J.S. Bullock, P. Burchat, D.L. Burke, G. Cagnoli, D. Calabrese, S. Chandrasekharan, S. Chesley, E.C. Cheu, J. Chiang, C.F. Claver, A.J. Connolly, K.H. Cook, A. Cooray, K.R. Covey, C. Cribbs, W. Cui, R. Cutri, G. Daubard, G. Daues, F. Delgado, S. Digel, P. Doherty, R. Dubois, G.P. Dubois-Felsmann, J. Durech, M. Eracleous, H. Ferguson, J. Frank, M. Freemon, E. Gangler, E. Gawiser, J.C. Geary, P. Gee, M. Geha, R.R. Gibson, D.K. Gilmore, T. Glanzman, I. Goodenow, W.J. Gressler, P. Gris, A. Guyonnet, P.A. Hascall, J. Haupt, F. Hernandez, C. Hogan, D. Huang, M.E. Huffer, W.R. Innes, S.H. Jacoby, B. Jain, J. Jee, J.G. Jernigan, D. Jevremovic, K. Johns, R.L. Jones, C. Juramy-Gilles, M. Juric, S.M. Kahn, J.S. Kalirai, N. Kallivayalil, B. Kalmbach, J.P. Kantor, M.M. Kasliwal, R. Kessler, D. Kirkby, L. Knox, I. Kotov, V.L. Krabbendam, S. Krughoff, P. Kubanek, J. Kuczewski, S. Kulkarni, R. Lambert, L. Le Guillou, D. Levine, M. Liang, K. Lim, C. Lintott, R.H. Lupton, A. Mahabal, P. Marshall, S. Marshall, M. May, R. McKercher, M. Migliore, M. Miller, D.J. Mills, D.G. Monet,

M. Moniez, D.R. Neill, J. Nief, A. Nomerotski, M. Nordby, P. O'Connor, J. Oliver, S.S. Olivier, K. Olsen, S. Ortiz, R.E. Owen, R. Pain, J.R. Peterson, C.E. Petry, F. Pierfederici, S. Pietrowicz, R. Pike, P.A. Pinto, R. Plante, S. Plate, P.A. Price, M. Prouza, V. Radeka, J. Rajagopal, A. Rasmussen, N. Regnault, S.T. Ridgway, S. Ritz, W. Rosing, C. Roucelle, M.R. Rumore, S. Russo, A. Saha, B. Sassolas, T.L. Schalk, R.H. Schindler, D.P. Schneider, G. Schumacher, J. Sebag, G.H. Sembroski, L.G. Seppala, I. Shipsey, N. Silvestri, J.A. Smith, R.C. Smith, M.A. Strauss, C.W. Stubbs, D. Sweeney, A. Szalay, P. Takacs, J.J. Thaler, R. Van Berg, D. Vanden Berk, K. Vetter, F. Virieux, B. Xin, L. Walkowicz, C.W. Walter, D.L. Wang, M. Warner, B. Willman, D. Wittman, S.C. Wolff, W.M. Wood-Vasey, P. Yoachim, H. Zhan und for the LSST Collaboration, *ArXiv e-prints* (2008).

[49] W. Jaffe, *Mon.Not.Roy.Astron.Soc.* **202** (1983), 995.

[50] I. Jee, E. Komatsu und S.H. Suyu, *JCAP* **1511** (2015), 033.

[51] I. Jee, E. Komatsu, S.H. Suyu und D. Huterer, *JCAP* **1604** (2016), 031.

[52] C.R. Keeton und L.A. Moustakas, *Astrophys. J.* **699** (2009), 1720.

[53] L. Knox, *Phys.Rev.* **D73** (2006), 023503.

[54] L. Koopmans, A. Bolton, T. Treu, O. Czoske, M. Auger et al. , *Astrophys.J.* **703** (2009), L51.

[55] L. Koopmans, T. Treu, C. Fassnacht, R. Blandford und G. Surpi, *Astrophys.J.* **599** (2003), 70.

[56] L.V.E. Koopmans, *Mon.Not.Roy.Astron.Soc.* **363** (2005), 1136.

[57] A. Kronawitter, R.P. Saglia, O. Gerhard und R. Bender, *Astron.Astrophys.* **144** (2000), 53.

[58] D.J. Lagattuta, S. Vegetti, C.D. Fassnacht, M.W. Auger, L.V.E. Koopmans und J.P. McKean, *Mon. Not. Roy. Astron. Soc.* **424** (2012), 2800.

[59] J. Larkin, M. Barczys, A. Krabbe, S. Adkins, T. Aliado, P. Amico, G. Brims, R. Campbell, J. Canfield, T. Gasaway, A. Honey, C. Iserlohe, C. Johnson, E. Kress, D. LaFreniere, J. Lyke, K. Magnone, N. Magnone, M. McElwain, J. Moon, A. Quirrenbach, G. Skulason, I. Song, M. Spencer, J. Weiss und S. Wright: *OSIRIS: a diffraction limited integral field spectrograph for Keck. OSIRIS: a diffraction limited integral field spectrograph for Keck*, In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Band 6269 von . (Juni 2006) Seite 62691A.

[60] N.M. Law et al. , *Publ. Astron. Soc. Pac.* **121** (2009), 1395.

[61] A. Lewis und S. Bridle, *Phys. Rev.* **D66** (2002), 103511.

[62] K. Liao, T. Treu, P. Marshall, C.D. Fassnacht, N. Rumbaugh et al. , *Astrophys.J.* **800** (2015), 11.

[63] E.V. Linder, *Phys.Rev.Lett.* **90** (2003), 091301.

[64] E.V. Linder, *Phys.Rev.* **D84** (2011), 123529.

[65] N. Lyskova, E. Churazov, I. Zhuravleva, T. Naab, L. Oser et al. , *Mon.Not.Roy.Astron.Soc.* **423** (2012), 1813.

[66] S.d. Mao und P. Schneider, *Mon. Not. Roy. Astron. Soc.* **295** (1998), 587.

[67] P.J. Marshall et al. , *Mon. Not. Roy. Astron. Soc.* **455** (2016), 1171.

[68] D. Merritt, *Astronomical.J.* **90** (1985), 1027.

[69] D. Merritt und L.A. Aguilar, *Mon.Not.Roy.Astron.Soc.* **217** (1985), 787.

[70] R.B. Metcalf und P. Madau, *Astrophys. J.* **563** (2001), 9.

[71] S. Miyazaki, Y. Komiyama, H. Nakaya, Y. Kamata, Y. Doi, T. Hamana, H. Karoji, H. Furusawa, S. Kawanomoto, T. Morokuma, Y. Ishizuka, K. Nariai, Y. Tanaka, F. Uraguchi, Y. Utsumi, Y. Obuchi, Y. Okura, M. Oguri, T. Takata, D. Tomono, T. Kurakami, K. Namikawa, T. Usuda, H. Yamanoi, T. Terai, H. Uekiyo, Y. Yamada, M. Koike, H. Aihara, Y. Fujimori, S. Mineo, H. Miyatake, N. Yasuda, J. Nishizawa, T. Saito, M. Tanaka, T. Uchida, N. Katayama, S.Y. Wang, H.Y. Chen, R. Lupton, C. Loomis, S. Bickerton, P. Price, J. Gunn, H. Suzuki, Y. Miyazaki, M. Muramatsu, K. Yamamoto, M. Endo, Y. Ezaki, N. Itoh, Y. Miwa, H. Yokota, T. Matsuda, R. Ebinuma und K. Takeshi: *Hyper Suprime-Cam. Hyper Suprime-Cam*, In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Band 8446 von *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. (September 2012) Seite 0.

[72] A. More, S.H. Suyu, M. Oguri, S. More und C.H. Lee, *Astrophys. J. Lett.* **835** (2017), L25.

[73] A. More et al. , *Mon. Not. Roy. Astron. Soc.* **456** (2016), 1595.

[74] S.T. Myers, C.D. Fassnacht, S.G. Djorgovski, R.D. Blandford, K. Matthews, G. Neugebauer, T.J. Pearson, A.C.S. Readhead, J.D. Smith, D.J. Thompson, D.S. Womble, I.W.A. Browne, P.N. Wilkinson, S. Nair, N. Jackson, I.A.G. Snellen, G.K. Miley, A.G. de Bruyn und R.T. Schilizzi, *Astrophys.J.* **447** (1995), L5.

[75] S.T. Myers et al. , *Mon. Not. Roy. Astron. Soc.* **341** (2003), 1.

[76] R. Narayan und M. Bartelmann: *Lectures on gravitational lensing. Lectures on gravitational lensing*, In *In Formation of Structure in the Universe. Cambridge Univ.* Press (1998).

[77] J.F. Navarro, C.S. Frenk und S.D.M. White, *Astrophys. J.* **490** (1997), 493.

[78] M. Oguri und P.J. Marshall, *Mon.Not.Roy.Astron.Soc.* **405** (2010), 2579.

[79] M. Oguri et al. , *Astron. J.* **132** (2006), 999.

[80] M. Oguri et al. , *Astron. J.* **135** (2008), 512.

[81] M. Oguri et al. , *Astron. J.* **143** (2012), 120.

[82] J.B. Oke et al. , *Publ. Astron. Soc. Pac.* **107** (1995), 375.

[83] L.P. Osipkov, *Soviet Astronomy Letters* **5** (1979), 42.

[84] D. Paraficz und J. Hjorth, *Astron. Astrophys.* **507** (2009), L49.

[85] R.M. Quimby, M. Oguri, A. More, S. More, T.J. Moriya, M.C. Werner, M. Tanaka, G. Folatelli, M.C. Bersten und K. Nomoto, *Science* **344** (2014), 396.

[86] S. Refsdal, *Mon.Not.Roy.Astron.Soc.* **128** (1964), 307.

[87] A.G. Riess et al. , *Astrophys. J.* **826** (2016), 56.

[88] C.E. Rusu, C.D. Fassnacht, D. Sluse, S. Hilbert, K.C. Wong, K.H. Huang, S.H. Suyu, T.E. Collett, P.J. Marshall, T. Treu und L.V.E. Koopmans, *ArXiv e-prints* (2016).

[89] P. Saha, C. Lobo, A. Iovino, D. Lazzati und G. Chincarini, *Astron. J.* **120** (2000), 1654.

[90] A.G. Sanchez, E.A. Kazin, F. Beutler, C.H. Chuang, A.J. Cuesta et al. , *Mon.Not.Roy.Astron.Soc.* **433** (2013), 1202.

[91] P. Schneider und D. Sluse, *Astron. Astrophys.* **559** (2013), A37.

[92] P. Schneider und D. Sluse, *Astron. Astrophys.* **564** (2014), A103.

[93] D. Sluse, J. Surdej, J. Claeskens, D. Hutsemekers, C. Jean et al. , *Astron.Astrophys.* **406** (2003), L43.

[94] A. Sonnenfeld, T. Treu, R. Gavazzi, S.H. Suyu, P.J. Marshall et al. , *Astrophys.J.* **777** (2013), 98.

[95] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J.F. Navarro, C.S. Frenk und S.D.M. White, *Mon. Not. Roy. Astron. Soc.* **391** (2008), 1685.

[96] M. Stiavelli und L.S. Sparke, *Astron.Astrophys.* **382** (1991), 466.

[97] S. Suyu, M. Auger, S. Hilbert, P. Marshall, M. Tewes et al. , *Astrophys.J.* **766** (2013), 70.

[98] S. Suyu, P. Marshall, M. Auger, S. Hilbert, R. Blandford et al. , *Astrophys.J.* **711** (2010), 201.

[99] S. Suyu, P. Marshall, R. Blandford, C. Fassnacht, L. Koopmans et al. , *Astrophys.J.* **691** (2009), 277.

[100] S. Suyu, T. Treu, S. Hilbert, A. Sonnenfeld, M. Auger et al. , *Astrophys.J.* **788** (2014), L35.

[101] S.H. Suyu et al. , *ArXiv e-prints* (2016).

[102] M. Tewes, F. Courbin, G. Meylan, C.S. Kochanek, E. Eulaers, N. Cantale, A.M. Mosquera, P. Magain, H. Van Winckel, D. Sluse, G. Cataldi, D. Vörös und S. Dye, *Astron.Astrophys.* **556** (2013), A22.

[103] T. Treu, A. Agnello und Strides Team: *The STRong-lensing Insights into Dark Energy Survey (STRIDES). The STRong-lensing Insights into Dark Energy Survey (STRIDES)*, In *American Astronomical Society Meeting Abstracts*, Band 225 von *American Astronomical Society Meeting Abstracts*. (Januar 2015) Seite 318.04.

[104] S. Unruh, P. Schneider und D. Sluse, *Astron. Astrophys.* (2016).

[105] S. Vegetti und L.V.E. Koopmans, *Mon.Not.Roy.Astron.Soc.* **392** (2009), 945.

[106] S. Vegetti und L.V.E. Koopmans, *Mon.Not.Roy.Astron.Soc.* **400** (2009), 1583.

[107] S. Vegetti, L.V.E. Koopmans, M.W. Auger, T. Treu und A.S. Bolton, *Mon. Not. Roy. Astron. Soc.* **442** (2014), 2017.

[108] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G.F. Snyder, D. Nelson und L. Hernquist, *Mon. Not. Roy. Astron. Soc.* **444** (2014), 1518.

[109] M.G. Walker, M. Mateo, E.W. Olszewski, J. Peñarrubia, N. Wyn Evans und G. Gilmore, *Astrophys.J.* **704** (2009), 1274.

[110] H. Witt, S. Mao und C. Keeton, *Astrophys.J.* **544** (2000), 98.

[111] J. Wolf, G.D. Martinez, J.S. Bullock, M. Kaplinghat, M. Geha, R.R. Muñoz, J.D. Simon und F.F. Avedo, *Mon.Not.Roy.Astron.Soc.* **406** (2010), 1220.

[112] S.A. Wright, J.E. Larkin, A.M. Moore, T. Do, L. Simard, M. Adamkovics, L. Armus, A.J. Barth, E. Barton, H. Boyce, J. Cooke, P. Cote, T. Davidge, B. Ellerbroek, A.M. Ghez, M.C. Liu, J.R. Lu, B.A. Macintosh, S. Mao, C. Marois, M. Schoeck, R. Suzuki, J.C. Tan, T. Treu, L. Wang und J. Weiss: *The infrared imaging spectrograph (IRIS) for TMT: overview of innovative science programs. The infrared imaging spectrograph (IRIS) for TMT: overview of innovative science programs*, In *Ground-based and Airborne Instrumentation for Astronomy V*, Band 9147 von . (Juli 2014) Seite 91479S.

[113] D. Xu, D. Sluse, L. Gao, J. Wang, C. Frenk, S. Mao, P. Schneider und V. Springel, *Mon. Not. Roy. Astron. Soc.* **447** (2015), 3189.

[114] D. Xu, D. Sluse, P. Schneider, V. Springel, M. Vogelsberger, D. Nelson und L. Hernquist, *Mon. Not. Roy. Astron. Soc.* (2015).

[115] H. Zhan, L. Knox und J.A. Tyson, *Astrophys. J.* **690** (2009), 923.

Acknowledgement

My Ph.D. has been a long journey, one which began in the United States and ended in Germany. I feel grateful and lucky that I could share my life with many people who have been part of it, who helped me, encouraged me and supported me throughout the process.
I cannot be more appreciative of my adviser, Eiichiro Komatsu, who kept me as his student the whole time. I appreciate his patience and understanding, for without him I would not be here. Sherry Suyu has been another adviser of mine, who helped me tremendously not only in terms of science, but also influenced me with her positive attitude and great personality. I would also like to thank Karl Gebhardt for having me as his student while I completed my masters program at the University of Texas at Austin.
I would like to thank my parents and brother who have always trusted me, encouraged me and supported me. I appreciate my two lovely cousins, Patricia Yunhee Jee and Heather Sunghee Jee who traveled a long way from Canada to visit me. I thank Yongyeon Choi who always encouraged me.
I enjoyed interacting with everybody who has been in the cosmology group over the years. I appreciate Donghui Jeong for being a great senior, for hosting me during my two visits to Baltimore, and for the enjoyable times with his family while he was visiting Munich.
I had fun with my bouldering friends. Among them I especially thank Titouan Lazeyras for being my music partner, Koki Kakiichi for being my climbing partner, and thanks again to both for being snacking buddies when work gets tiring.
I acknowledge my friends who kept in touch with me, and have been enormously supportive. I thank Yewon Choi, Hyesoo Choi, Amy Jones, Durand D'souza, Haakon Andresen, Marius Eide, Malin Renneby, Eliceth Rojas, Yoojin Jang, and Joyce Jiyoung Whang.
Having been at two institutes, I had many great office mates: Talking to Andrew Riddle and Roberto Hernandes at UT Austin was always relaxing. It was fun and encouraging to have hard working office mates like Anne Klitsch and Dani Chao at MPA.
Living in a country where I didn't speak the language was another challenge. I thank Philipp Wullstein, Christian Wagner and Fabian Schmidt for helping me go through German documents. I appreciate Akın Yıldırım for translating the abstract of this thesis into German.
At last but not least, I would like to send my best appreciation to Mike Anderson, with whom I went through both good and hard times together, who has always been supportive and patient in every aspect of my life in Munich.